THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Characterization of the interactome of the porcine reproductive and respiratory syndrome virus (PRRSV) NSP2 protein reveals the hyper variable region as a binding platform for association with 14-3-3 proteins

Citation for published version:
Xiao, Y, Wu, W, Gao, J, Smith, N, Burkard, C, Xia, D, Zhang, M, Wang, C, Archibald, A, Digard, P, Zhou, E & Hiscox, JA 2016, 'Characterization of the interactome of the porcine reproductive and respiratory syndrome virus (PRRSV) NSP2 protein reveals the hyper variable region as a binding platform for association with 14-3-3 proteins', *Journal Of Proteome Research*, vol. 15, no. 5, pp. 1388-1401. https://doi.org/10.1021/acs.jproteome.5b00396

Digital Object Identifier (DOI):
10.1021/acs.jproteome.5b00396

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal Of Proteome Research

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

Characterization of the interactome of the porcine reproductive and respiratory syndrome virus (PRRSV) nsp2 protein reveals the hyper variable region as a binding platform for association with 14-3-3 proteins.

Yihong Xiao[†,‡], Weining Wu[†], Jiming Gao[§], Nikki Smith[‖], Christine Burkard[‖], Dong Xia[†], Minxia Zhang[‡], Chengbao Wang[§], Alan Archibald[‖], Paul Digard[‖], En-min Zhou[§] and Julian A. Hiscox[*,†,§] .

[†]Department of Infection Biology, Institute of Infection and Global Health, University of Liverpool, Liverpool, L3 5RF, UK.

[‡]Department of Basic Veterinary Medicine, College of Animal Science and Veterinary Medicine, Shandong Agricultural University, Tai'an, P.R. China, 271018.

[§]College of Veterinary Medicine, Northwest A&F University, Yangling, P.R. China, 712100.

[‖]The Roslin Institute, University of Edinburgh, Edinburgh, UK.

[*]Corresponding author: Email: julian.hiscox@liverpool.ac.uk; Tel: +44 (0)151 795 0222.

Abstract

Porcine reproductive and respiratory syndrome virus (PRRSV) is a major threat to the swine industry worldwide and hence global food security, exacerbated by a newly emerged highly pathogenic (HP-PRRSV) strain from China. PRRSV non-structural protein 2 (nsp2) is a multifunctional polypeptide with strain-dependent influences on pathogenicity. A number of discrete functional regions have been identified on the protein. Quantitative label free proteomics was used to identify cellular binding partners of nsp2 expressed by HP-PRRSV. This allowed the identification of potential cellular interacting partners and the discrimination of non-specific interactions. The interactome data was further investigated and validated using biological replicates and also compared with nsp2 from a low pathogenic (LP) strain of PRRSV. Validation included both forward and reverse pulldowns and confocal microscopy. The data indicated that nsp2 interacted with a number of cellular proteins including; 14-3-3, CD2AP and other components of cellular aggresomes. The hyper variable region of nsp2 protein was identified as a binding platform for association with 14-3-3 proteins.

Keywords

Porcine reproductive and respiratory syndrome virus, non-structural protein 2, label free proteomics, proteomics, virus, interactome, aggresomes.

Introduction

Approximately 25 years ago, a disease called “mystery swine disease” characterized by high fever, disordered respiratory, abortion, and stillborn or mummified piglets emerged into Northern American and in Central Europe.[1] The etiologic agent was identified as a positive sense, single-stranded RNA virus,[2, 3] designated porcine reproductive and respiratory syndrome virus (PRRSV); subsequently classified into the order *Nidovirales*, family *Arteriviridae*, genus.[4] After further genomic analysis, two major genotypes were identified; the European (type 1) and the North American (type 2) isolates. This disease is now found worldwide and causes significant losses in the swine industry owing to reproductive disorders and growth retardation. PRRSV is prone to mutate and recombine *in vivo* resulting in new strains, complicating its control and eradication.[5-7] The emergence of a highly pathogenic strain of PRRSV (HP-PRRSV) caused devastating economic losses in China in 2006.[8, 9]

The genome of PRRSV is approximately 15kb in length and contains ten open reading frames (ORFs) in succession: ORF1a-ORF1b-ORF2a-ORF2b-ORF3-ORF4-ORF5-ORF5a-ORF6-ORF7.[10] Type 1 and type 2 genotypes are highly diverse, sharing only approximately 60% nucleotide identity.[11, 12] ORF1a and ORF1b occupy three quarters of the whole genome and encode two large replicase precursor polyproteins pp1a and pp1ab. These polyproteins are then processed into 13–16 functional non-structural proteins (nsps) by viral proteinases directed by ORF1a .[13, 14]

Of these mature polypeptides, nsp2 is the largest cleavage product and the most variable in terms of length and amino acid sequence with less than 40%

shared at the amino acid level between both genotypes[11], due in part to intragenic and intergenic recombination.[6, 7] These differences contribute to virulence and genotyping of PRRSV.[8,9,11,12] Nsp2 possesses a cysteine protease (PL2) domain, a central hypervariable domain (HV), a trans-membrane domain (TM) followed by a tail domain.[14, 15] The PL2 domain possesses dual characteristics of papain-like cysteine proteases and chymotrypsin-like serine proteases and is responsible for cleavage of nsp2/3 as well as acting as a co-factor of the nsp4 serine protease.[10, 16, 17] The PL2 region is also a member of the ovarian tumour (OTU) protease super-family. This family of proteases is thought to be involved in the inhibition of the host innate immune response.[18] Supporting this for PRRSV, the PRRSV OTU domain has been shown to antagonize the induction of type I interferon and is capable of deconjugating ubiquitin and ISG15.[19]

The HV region located between aa160-844 contributes to the size differences observed between different strains and types of PRRSV.[10-12] Mutations, insertions, and more commonly, deletions are observed in this region. The genetic marker of HP-PRRSV is the deletion of 30 aa in the central part of nsp2. Analysis has indicated that the central part can tolerate nearly 400aa (324-726) deletion but cannot be removed as a whole.[20] The C-terminal region of nsp2 contains three to four TM domains with unknown function, although these may be involved in tethering replication complexes. At least 6 isoforms of nsp2 have been detected in PRRSV infected cells; all of them were stable and had low turnover rates.[21] Despite its name, nsp2 is found in virions, suggesting that this protein may also be a virion-associated structural protein.[22] Nsp2 also contains predominant B cell epitopes and

evokes high titer antibody resulting in modulating adaptive immune response.[23]

Many viral proteins including components of RNA dependent RNA polymerase complexes form associations with cellular proteins inside virus infected cells. This can reflect a balance between pro-viral and anti-viral activity. Given the complexity and multi-functional nature of the nsp2 protein we hypothesized that it would interact with multiple cellular proteins. To test this, label free quantitative proteomics and other techniques were used to identify cellular proteins that potentially formed specific interactions with GFP-tagged nsp2 from HP-PRRSV as well as classical low pathogenicity strains (LP-PRRSV). We identified 91 cellular proteins as potential partners of nsp2 from HP-PRRSV, including many that were part of larger complexes, such as the aggresome. Notably, this included the 14-3-3 protein family, which plays a role in critical regulatory processes in the cell.

Experimental Procedures

Cells and virus

293T, MARC-145 and PK-15 cells were obtained from the Health Protection Agency Culture Collections and cultured in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 10% fetal bovine serum at 37°C, 5% CO_2 in a humidified incubator. PRRSV strain of the classical low pathogenicity CH-1R and high pathogencity strain SD-16 with 30-aa deletion in nsp2 gene (isolated from a clinical case in 2007, the genome sequence was submitted in GenBank as NO JX087437) were propagated in MARC-145 cells. Viruses were harvested 3 days post-infection and passed through 0.22 μm filter. Virus titre was measured by tissue culture 50% infectious dose and calculated by the Reed-Muench method[24] depending on observation of CPE. HEK293FT (Life Technologies) and MARC-145 (ATCC) were used to express and propagate PRRSV-nsp2-GFP (North American type 1 virus SD01-08),[25] respectively.

PRRSV-nsp2-GFP expression and propagation

HEK293FT cells at about 80% confluency, propagated the day before, were transfected with 8μg pSD01-08-GFP plasmid 24 per T75 flask using polyethylenimine (PEI, linear, 25,000kDa, Alfa Aesar; 1μg/μl stock concentration) transfection reagent at 1:5 ratio. Medium was replaced 4h post transfection by infection medium (2% FBS, 100U/ml Penicillin and 100μg/ml Streptomycin (Gibco) in DMEM). Supernatant was harvested 72h

post transfection, filtered through a 0.22 μm filter, and buffered with 1M HEPES (pH 7.2-7.5, Gibco) to 25mM final concentration before freezing. Highly confluent MARC-145 cells were infected with supernatant from HEK293FT cells at a 1:10 dilution (approx. MOI=0.0001) in infection medium. Inoculum was replaced 2 hours post infection (hpi) and infection allowed to continue until 60hpi. Supernatant was harvested and buffered to 25mM HEPES as described above. Virus was passaged two more times at MOI=0.1 on MARC-145 cells to yield cell adapted virus. Infections and harvests were performed as described above, yielding a passage three PRRSV-nsp2-GFP stock at 9.3 x 10^6 $TCID_{50}$/ml. Virus infectivity was assessed by endpoint dilution on MARC-145 cells. Confluent MARC-145 cells were infected with serial dilutions of viral stocks (after 1x freezing cycle) in infection medium. Inoculum was replaced 2hpi by infection medium. Infection was assessed 48hpi and confirmed 72hpi by visualization of GFP expression using fluorescence wide-field microscopy.

Plasmids

Nsp2 genes were amplified from recombinant plasmids (constructed by others in our group) containing the full length of ORF1a by the following primers with *Bgl* II and *Sma* I (underline) restriction sites: 5'-gatcagatctgctggaaagagagcaaggaaagc-3', 5'-tccccgggccttatcccgaaggcttgg aaatttgcc-3'. The amplified products (predicted to be 3582bp (1194aa) for the nsp2 gene from LP-PRRSV strain CH-1R and 3492bp (1164aa) from

HP-PRRSV strain SD-16) were ligated into the pEGFP-C1 vector to produce the recombinant plasmids EGFP-LPnsp2 and EGFP-HPnsp2.

Transfection

293T cells were seeded and were 50% confluent prior to transfection. For 10 cm dishes 10 μg of recombinant plasmids and 61μl of $CaCl_2$ (2M) were diluted into ddH_2O to make 500 μl mixture. The mixtures were added into the same volume of 2X HBS (50mM HEPES, 280mM NaCl, 1.5mM Na_2HPO_4, pH 7) and incubated for 30 minutes at room temperature. The final mixtures were added dropwise to the cells with the volume of 1 ml. Then the cells were grown at 37°C, 5% CO_2 for 24 hours. For 12-well plates, 0.71 μg of DNA and 4.37μl of 2 M $CaCl_2$ were mixed with 36 μl 2X HBS instead. A total of 72 μl mixture was added to the cells. Monolayer MARC-145 and PK-15 cells were transfected with Lipofectamine 2000 (Sigma) according to the instruction.

GFP pull downs

293T cells were transfected with plasmids EGFP-LPnsp2 and EGFP-HPnsp2 using the calcium phosphate method, using four 10 cm dishes per plasmid. Cells were harvested 24 h post-transfection in lysis buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5% (v/v) NP40) with halt protease inhibitor (Fisher Scientific) at a final concentration of 1X. Clarified cell lysates were incubated with GFP-trap beads (Chromo Tek) for 2 h at 4°C before three washes with 10 mM Tris HCl, pH 8, 0.1% NP-40, 150 mM NaCl, 0.5 mM EDTA. For LC-MS/MS analysis, the bound proteins were eluted with 100 μl of glycine elution buffer (0.2 M glycine, pH 2.5), the supernatant was

collected and neutralized with 10 μl of 1M Tris base (pH 10.4). Samples were prepared separately in triplicate. For western blotting analysis, bound proteins were eluted with 100 μl 2x SDS-Sample buffer. RFP-Trap coupled agarose beads were used as a negative control.

Sample preparation for proteomics

The glycine buffer eluted samples were diluted two fold with 25 mM ammonium bicarbonate. Rapigest (Waters) was added to a final concentration of 0.05% (w/v) and the sample was heated at 80°C for 10min. Proteins were reduced with 3 mM dithiothreitol (Sigma) at 60°C for 10min then alkylated with 9 M iodoacetimide (Sigma) at room temperature for 30min in the dark. Proteomic grade trypsin (0.2μg, Sigma) was added and samples were incubated at 37°C overnight. The Rapigest was removed by adding TFA to a final concentration of 1% (v/v) and incubating at 37°C for 2h. Peptide samples were centrifuged at 12,000g, 4°C for 1h to remove precipitated Rapigest. Each digest was concentrated and desalted using C18 Stage tips (Thermo scientific), then dried down using a centrifugal vacuum concentrator (Jouan) and resuspended in a 0.1% (v/v) TFA, 3% (v/v) acetonitrile solution.

NanoLCMSESI MS/MS analysis

Peptide mixtures (2μl) were analyzed by on-line nanoflow liquid chromatography using the nanoACQUITY-nLC system (Waters MS technologies, Manchester, UK) coupled to an LTQ-OrbitrapVelos (ThermoFisher Scientific) mass spectrometer equipped with the manufacturer's nanospray ion source. The analytical column (nanoACQUITY

UPLCTM BEH130 C18 15cm x 75μm, 1.7μm capillary column) was maintained at 35°C and a flow-rate of 300nl/min. The gradient consisted of 3-40% acetonitrile in 0.1% formic acid for 50 min then a ramp of 40-85% acetonitrile in 0.1% formic acid for 3 min. Full scan MS spectra (m/z range 300-2000) were acquired by the Orbitrap at a resolution of 30,000 with precursor isolation width of 1.2 m/z. Analysis was performed in data dependent mode. The top 20 most intense ions from MS1 scan (full MS) were selected for tandem MS by collision induced dissociation (CID) at normalized CE 35%, and all product spectra were acquired in the LTQ ion trap. Charge state rejection was enabled where charge state one was rejected and dynamic exclusion was enabled at repeat count one, repeat duration 20 and exclusion duration 20.

Thermo RAW files were imported into Progenesis LC–MS (version 4.1, Nonlinear Dynamics). Runs were time aligned using default settings and using anauto selected run as reference. Peaks were picked by the software and filtered to include only peaks with a charge state of between +2 and +6. Peptide intensities were normalised against the reference run by progenesis LC-MS and these intensities are used to highlight differences in protein expression between GFP control and GFP-nsp2 samples with supporting statistical analysis (ANOVA p-values) calculated by the Progenesis LC-MS software. MS. Spectral data were transformed to mgf files with Progenesis LC–MS and exported for peptide identification using the locally installed PEAKS Studio 7 (Bioinformatics Solutions Inc.) search engine. Multiple search engine platform provided by PEAKS Studio named inChorus was used, which combines searching results from PEAKS DB (Bioinformatics Solutions

Inc.), locally installed Mascot (v2.3, Matrix Science) and OMSSA (National Center for Biotechnology Information) and X!Tandem (Global Proteome Machine Organization) that were bundled with PEAKS Studio 7. Tandem MS data were searched against the human and PRRSV predicted proteomes custom database that contained the common contamination (The Global Proteome Machine, version 2012.01.01, 115 sequences) and internal standards (MassPREP Digestion Standard Mix 2, Waters, 4 sequences), host cell proteome sequences (UniProt reviewed, version February 2014, 20,276 sequences) and viral sequences (GenBank JX087437). Search parameters were as follows; enzyme set to trypsin, precursor mass tolerance set to 10 ppm and fragment mass tolerance set to 0.5 kDa. One missed tryptic cleavage was permitted. Carbamidomethylation (cysteine) was set as a fixed modification and oxidation (methionine) set as a variable modification. The false discovery rates were set at 1% and at least two unique peptides were required for reporting protein identifications. Results were imported into Progenesis LC–MS as .xml files

Immunoprecipitation (IP)

293T cells were transfected with the plasmids using calcium phosphate. The cell lysates were prepared and incubated with 2 μg of anti-14-3-3 (Abcam, ab6081, rabbit polyclonal antibody), anti-DNAJA2 (Santa Cruz, sc-136515, mouse monoclonal antibody), anti-nucleolin (Abcam, ab13541 mouse monoclonal antibody) anti-CD2AP(Cell Signaling, A599, rabbit polyclonal antibody), anti-HSP70 (Sigma, H5147, mouse monoclonal antibody), or anti-ribosomal protein S6 (FL-249) (Santa Cruz, sc-20085, rabbit polyclonal

antibody) antibodies overnight on a rotator at 4°C before precipitation by the addition of 50 μl protein G resin (Gneron) for a further 2h. After washing, bound proteins were eluted with 100 μl 2x SDS-Sample buffer.

Western blotting

Samples were separated by 10-15% SDS-PAGE and transferred to PVDF membranes (Millipore) using a Bio-Rad semi-dry transfer apparatus according to standard procedures. The primary antibodies were anti-GFP (Santa Cruz, sc8334) and all the antibodies mentioned in IP method. HRP-conjugated anti-mouse or anti-rabbit secondary antibodies were purchased from sigma. Blots were imaged by using Clarity™ Western ECL subtract (BIO-RAD).

Aggresome detection and microscopy

293T, MARC-145 and PK-15 cells were seeded onto coverslips and transfected with plasmids pEGFP-C1, EGFP-LPnsp2 and EGFP-HPnsp2 by Lipofectamine 2000 (Invitrogen), respectively. MARC-145 cells were infected with PRRSV strains at 0.01 MOI. Twenty-four hours post-transfection or post-infection the cells were fixed with 4% paraformaldehyde and permeabilised with 0.1% (v/v) triton X-100 in phosphate buffered saline. For aggresome detection, cells were incubated with aggresome detection reagent and Hoechst 33342 supplied by Aggresome Detection Kit (Abcam, ab139486). Transfected 293T and MARC-145 cells were probed with anti-14-3-3, anti-CD2AP, anti-HSP70 (mentioned above) and anti-vimentin (Abcam, ab92547, rabbit monoclonal antibody) antibodies. Transfected PK-15 cells were probed with anti-14-3-3 and anti-CD2AP antibodies. The proteins were

visualized with either Alexa Fluor®568 rabbit anti-mouse or Alexa Fluor® 546 donkey anti-rabbit antibodies (Life Technologies). PRRSV infected MARC-145 cells were incubated with anti-HP-PRRSV SD16 pig hyper immunized serum which was prepared by artificial infection of HP-PRRSV SD16 strain under lab condition, and visualized with FITC-goat anti-pig IgG and 14-3-3 protein was visualized by Cy3-goat anti-rabbit IgG (Jackson). All probed cells were observed under fluorescence microscope (Carl Zeiss Upright Axio Lab.A1).

Analysis of soluble and insoluble fractions

293T cells were transfected with pEGFP-C1, EGFP-LPnsp2 and EGFP-HPnsp2. Twenty-four hours post-transfection, cells were collected by cell scraper and washed three times with PBS. The cells were then lysed in lysis buffer with halt protease inhibitor for 30 min on ice before clarification by centrifugation at 1,000×g, 3 min to remove cell debris. The supernatant was collected and centrifuged at 20,000×g for 30 min and both the supernatant and pellet saved. The pellet was washed twice with RIPA buffer (NP40:1%, 350 mM Tris-HCL, 150 mM NaCl, 0.1% SDS). The insoluble parts were collected by centrifugation with 20,000 ×g for 1 h and resolved in 8M urea lysis buffer (Urea 8M, DTT 1M, Tris 1.5M, 0.1% SDS, Glycerol 10%,). The same volumes were loaded onto SDS-PAGE for western blot analysis.

Results

Expression of EGFP-HPnsp2 and EGFP-LPnsp2 in 293-T cells and interaction analysis.

To determine the interactome of the HP-PRRSV nsp2 protein, a combined EGFP-trap LC-MS/MS approach was used, similar to the approach we have used to elucidate the interactomes of other viral proteins, including human respiratory syncytial virus (HRSV) NS1 and L proteins , infectious bronchitis virus and PRRSV nucleocapsid proteins, and Ebola virus VP24.[26-30] The nsp2 gene of HP-PRRSV was cloned downstream of the EGFP gene to direct expression of a fusion protein consisting of EGFP N-terminal to nsp2 (EGFP-HPnsp2). Cell lysates were prepared from cells expressing EGFP-HPnsp2 or EGFP alone, and the GFP proteins purified by EGFP-trap chromatography to pull down associated cellular proteins. Label free LC-MS/MS was used to identify and quantify the abundance of these with the expectation that specific interaction partners of nsp2 would be selectively enriched in the EGFP-HPnsp2 samples.

Despite PRRSV infecting swine and naturally replicating in porcine alveolar macrophages (PAMs), human 293T cells were used to derive potential interactomes for several reasons. (1). 293T cells are highly susceptible to calcium phosphate transfection of plasmids allowing almost ubiquitous high expression of the recombinant protein for analysis. (2). Protein identification can be referenced to human databases that are more complete that swine databases. (3). Many studies on PRRSV use the MARC-145 simian cell line as one of the few cell lines that reliably supports PRRSV replication. (4). PAMS are refractive to expression of exogenous

proteins. (5). It is hard to justify the number of PAMs required for biochemical experiments under the replacement, reduction and refinement in animal research, when other model systems (such as 293T cells can be used).

The expression of EGFP, EGFP-HPnsp2 and EGFP-LPnsp2 were confirmed in 293 T cells by both direct fluorescence (Figure 1A) and western blot (Figure 1B). All three proteins were expressed in similar numbers of cells, while western blot analysis showed that the full-length fusion proteins were the expected molecular weight (Figure 1B). EGFP-LPnsp2 expression also produced several lower molecular weight products (Figure 1B), possibly associated with cleavage of the protein. A highly specific GFP-trap was used to selectively precipitate EGFP and the EGFP fusion proteins. Pull down products from these trap experiments were digested with trypsin after adding RapiGest to optimize in-solution protein digestion. The peptides were desalted and enriched for analyzing by label-free quantitative proteomics.

To provide a statistically robust dataset, pull downs with both the EGFP control and EGFP-HPnsp2 were performed independently in duplicate. The abundance of bound peptides from the EGFP and EGFP-HPnsp2 traps were compared by Progenesis LC–MS software against a human database. The use of quantitative label free proteomics allowed assignment of proteins that bound to either EGFP and/or the binding matrix compared to proteins that bound to the EGFP-HPnsp2. Selection criteria for potential interacting partners for HPnsp2 included a False Discovery Rate (FDR) ≤1% two or more unique peptides for reporting protein identifications and a binding ratio greater than 2-fold.

The identified proteins from each experiment are combined and shown in Table One. These proteins were analysed using the STRING algorithm in order to investigate potential interactions. After enrichment of KEGG Pathways, these proteins were divided into 4 categories: the distinct ribosomal complex cluster composed of 40 ribosomal proteins, the smaller cluster composed by 5 isoforms of 14-3-3 protein which play an important role in signaling pathways, four proteins (SSRG, SSRD, DNAJA2 and PDIA3) involved in protein processing in endoplasmic reticulum and unassigned proteins such as CTBP2, CD2AP, and NUCL. Analysis using Ingenuity Pathway Analysis highlighted several proteins which could be clustered into similar functions including those involved in gene expression (EIF3I, RBM3, RBM4, CDK6, CTPBP2 and PHB2), protein synthesis (EIF3I, RBM3, RBM4 and PDIA3) and apoptosis (CD2AP, CDK6, COX5A, CTBP2, EIF3I, PDIA3, PHB2 and RBM3).

Validation of the EGFP-HPnsp2 interactome.

The LC-MS/MS identified several different sets of proteins, so selected interactions were validated in repeat GFP trap pull downs, but using antibodies to specific proteins to confirm interactions using western blot (Figure 2A). As further controls of specificity, EGFP-LPnsp2 was also tested as a bait polypeptide, while negative controls were no EGFP bait, EGFP only as bait, or using RFP trap beads (which would not be expected to enrich the bait polypeptides) in the presence of the nsp2 fusion proteins. Cellular target proteins were selected on the basis of variability in peptides used for identification, difference in ratio between EGFP and EGFP-HPnsp2 and bioinformatic analysis of function. This led to the choice of RS6 (3 peptides,

13.5-fold enhanced) from the ribosomal protein cluster, 14-3-3 (5 peptides, 6.7-fold enhanced) from the 14-3-3 family cluster, DNAJA2 (2 peptides, 4.71-fold enhanced) in the protein processing in endoplasmic reticulum cluster and selected examples of unassigned proteins: CD2AP (12 peptides, 11.7-fold enhanced) and NUCL (23 peptides, 5. 6-fold enhanced). HSP70 (4 peptides, 1.5-fold enhanced) was also investigated. Blotting for GFP showed that as expected, the GFP bait polypeptides were efficiently pulled down by the GFP trap beads, but not the RFP trap beads (Figure 2A, top two rows). Unambiguous specific binding of nucleolin, CD2AP, HSP70, and DNAJA2 to both HP and LP nsp2 polypeptides but none of the negative controls was observed. Evidence of similar binding of the ribosomal RS6 protein was also seen, but the poor quality of the available antibody lessened the signal-to-noise ratio of the data. Clear binding of 14-3-3 to the EGFP-HPnsp2 protein was also seen, but not to the counterpart LP fusion protein, whose binding to this protein was no different to that of the EGFP only control (Figure 2A).

A reverse pull down approach was also used to confirm interactions, in which antibodies to the specific cellular proteins were tested for their ability to precipitate the EGFP-tagged nsp2 proteins (Figure 2B). The binding of these proteins to either EGFP-HPnsp2 or EGFP-LPnsp2 was compared to an EGFP control. The data indicated that both EGFP-HPnsp2 and EGFP-LPnsp2, but not EGFP, associated with NUCL, CD2AP, HSP70, DNAJA2, and 14-3-3. RS6 protein precipitated EGFP-HPnsp2 but EGFP-LPnsp2 was either below the limit of detection or did not bind to it. Thus overall, the interaction of each of the chosen cellular proteins with both HP and LP PRRSV nsp2 was

validated by specific biochemical assays, although potentially, some strain-dependent variation was seen in the 14-3-3 and RS6 interactions.

Confocal microscopy was then used as an independent approach to explore potential interactions between either EGFP-HPnsp2 or EGFP-LPnsp2 and the cellular proteins (where antibody combinations allowed), again using EGFP as a control (Figure 3). These three exogenous proteins were over expressed separately in 293T cells (used for the pull down), MARC-145 cells (a permissive cell line for PRRSV) or PK-15 cells (a porcine derived cell line). The data indicated that for both cell lines EGFP-HPnsp2 and EGFP-LPnsp2 potentially interacted with 14-3-3 and CD2AP and shared similar locations in the cell. These two cellular proteins did not co-localise with EGFP. The co-localization of nsp2 protein with 14-3-3 and CD2AP in PK-15 cells, further confirmed that nsp2 could interact with 14-3-3 and CD2AP There was no evidence for co-localization between either EGFP-HPnsp2 or EGFP-LPnsp2 and RS6.

Interactions between nsp2 and cellular proteins occur in virus-infected cells. The interactome data for nsp2 was generated in the context of over-expression of the isolated viral polypeptide. To investigate whether these interactions occurred in PRRSV infected cells where nsp2 is expressed in a native context in the presence of other viral proteins we made use of a recombinant LP-PRRSV (Type 1 genotype) that expressed nsp2 as a biologically active EGFP fusion protein (termed in this study PRRSVnsp2-EGFP). MARC-145 cells were either mock-infected, infected with PRRSVnsp2-EGFP or, to provide a negative control, transfected with EGFP, and cell lysates prepared for GFP trap pull downs (Figure 4A).

Precipitation of specific cellular protein targets was investigated by western blot. Examination of the input cell lysates showed that antibodies raised against human isoforms of HSP70, 14-3-3, NUCL and CD2AP the equivalent polypeptides from the MARC-145 cell line. None of these proteins interacted detectably with the EGFP control nor were present in the bound mock-infected sample (Figure 4A). In contrast, HSP70, 14-3-3 and NUCL were all trapped by nsp2-EGFP expressed in the context of recombinant PRRSV infection. However, CD2AP, which associated with overexpressed EGFP-HPnsp2 or EGFP-LPnsp2, either did not associate with EGFP-nsp2 in the viral context or did so below the limits of detection. The interaction of 14-3-3 with PRRSV was also investigated in MARC-145 cells infected with either LP-PRRSV or HP-PRRSV (Figure 4B). The data indicated that 14-4-3 accumulated at the sites of virus replication and co-localized with viral proteins (Figure 4B). This may also resemble the re-localization of 14-3-3 observed in some cases with over-expression of nsp2.

Nsp2 associates with cellular aggresomes.

The interactome analysis indicated that both HPnsp2 and LPnsp2 formed interactions with a number of proteins involved in the formation of aggresomes, [31] including chaperone and 14-3-3 proteins. Whilst aggresomes are part of the misfolded protein response, a number of viruses initiate and use and form these complexes to facilitate virus replication.[32,33] For example, the cytomegalovirus UL76 protein initiates aggresome formation.[34] To investigate whether aggresomes formed in the presence of nsp2 we used complementary imaging and biochemical approaches to examine the aggresome complex. First, to identify the aggresome complex in either cells

expressing either EGFP-HPnsp2, EGFP-LPnsp2 or EGFP (as a control), cells were treated with a dye that becomes fluorescent when bound to aggregated cargo[35] and imaged by confocal microscopy (Figure 5A). The data indicated that in both 293T and MARC-145 cells, both EGFP-HPnsp2 and EGFP-LPnsp2 co-localized with the marker for the aggresome (as determined by the yellow signal in the merged image) (Figure 5A). To investigate this further, two components of the aggresome, HSP70 and vimentin, were visualized in 293T and MARC-145 cells expressing EGFP-HPnsp2, EGFP-LPnsp2 or EGFP (as a control) (Figure 5B). The data indicated that EGFP-HPnsp2 and EGFP-LPnsp2 could co-localize with either HSP70 or vimentin, whereas EGFP did not (Figure 5B).

Recently, 14-3-3 has been reported to target misfolded chaperone-associated proteins to aggresomes by acting as an adaptor along with an HSP70 co-chaperone, Bcl-2-associated athanogene 3 (BAG3) to recruit chaperone-associated protein cargos to dynein motors for their transport to aggresomes.[36] In the GFP-nsp2 pull down, seven different isoforms of 14-3-3, HSP70 and DNAJA2 were identified as potentially interacting with nsp2. We therefore hypothesised that 14-3-3 might play a key role in the formation of chaperone-dependent nsp2 aggresomes. To test this, we isolated detergent-insoluble cell fractions[37] expected to be enriched for aggresomes from 293T cells transfected with the EGFP polypeptides and probed the distribution of EGFP, CD2AP, HSP70, DNAJ2, 14-3-3 and (as a control), GAPDH (Figure 6). As expected, GAPDH was not detectable in the insoluble fractions, indicating the reliability of this isolation method. EGFP-HPnsp2 and EGFP-LPnsp2 were detected in both the soluble and the

insoluble fraction, whereas EGFP was only detected in the soluble fraction, despite its higher expression level (Figure 6). This is consistent with the nsp2 polypeptide contributing to aggresome formation. A fraction of HSP70 was present in the insoluble fraction in EGFP-transfected cell lysates, but the proportion increased in the nsp2-transfected samples. Furthermore, CD2AP, DNAJA2, different 14-3-3 isoforms were only identified in the insoluble fraction in the presence of the EGFPnsp2 polypeptides. This result is consistent with the hypothesis that CD2AP, DNAJA2 and HSP70 as well as 14-3-3 polypeptides participate in nsp2 aggresome formation.

To determine whether CD2AP, DNAJA2 and HSP70 interact with 14-3-3, a reverse pull down was performed using the 14-3-3 antibody (Figure 6). CD2AP was not detected in the pull down product while both DNAJA2 and HSP70 were detected, indicating 14-3-3 only interacted with DNAJA2 and HSP70 but not CD2AP. 14-3-3 and HSP70 interacted with both EGFP and EGFP-HPnsp2 while DNAJA2 interacted with EGFP-HPnsp2.

Specific domains of nsp2 interact with cellular protein targets.

The interactome data indicated that nsp2 interacts with multiple protein targets; suggesting this protein may act as a hub protein, behavior sometimes expected for conformationally versatile proteins with disordered regions 37. A protein disorder prediction algorithm did indeed suggest the existence of disordered regions within the HPnsp2 polypeptide, with the most disordered region corresponding to the HV region (amino acids 160-814) (Figure 7). To test whether different domains of HPnsp2 interacted with specific cellular proteins, the the OTU, HV, TM and Tail regions were expressed as separate EGFP fusion proteins (Figure 8A). Immunofluorescence (Figure 8B) and

western blot (Figure 8C) analysis indicated that all fragments of nsp2 were successfully expressed in 293T cells, although the TM construct accumulated to noticeably lower levels than the rest.

The ability of the individual nsp2 domains to bind nucleolin, CD2AP, HSP70, DNAJ2 and 14-3-3 was then tested by GFP trap analysis and western blot as before. (Figure 9A). As a further control to assess specificity, the association of the EGFP-HPnsp2 polypeptides with HSP90 was also examined. HSP90 was not identified in the interactome analysis as binding to EGFP-HPnsp2, and therefore would not be predicted to interact with full length or individual domains of nsp2. Western blot analysis confirmed this prediction (Figure 9A). The EGFP control also failed to precipitate any of the cellular targets, as expected. In contrast, all five potential nsp2 interaction partners were precipitated by the full length EGFP-HPnsp2 polypeptide. HSP70 bound to all subfragments of nsp2 other than the TM domain, although the poor expression of the latter construct prevented firm conclusions about its behaviour. DNAJA2 did not bind to any of the individual nsp2 domains, while nucleolin, CD2AP and 14-3-3 all bound specifically to the HV region. Thus the HPnsp2 HV region functions to bind multiple cellular polypeptides, consistent with the concept that it acts as a binding hub.

Discussion

Many viruses such as PRRSV synthesize large polyproteins that are then cleaved into functional sub-units. For PRRSV and related arteriviruses and coronaviruses these proteins are involved in the replication and transcription of the viral genome, virus biology and modulation of host cell biology. Here we investigated the interaction of the nsp2 protein encoded by the PRRSV ORF1a gene as part of a larger polyprotein. The precise role of nsp2 in virus biology has not been defined and recent data has shown the protein is also incorporated in virus particles.[22] To investigate the cellular interactome of nsp2, proteins were expressed from low and highly pathogenic variants of PRRSV (Figure 1). Several LPnsp2 isoforms were detected, while there was one specific species associated wih the expression of HPnsp2 (Figure 1B). Therefore, in this work we focused on the interaction between HPnsp2 and the host cell, by first determining the cellular interactome of the nsp2 protein expressed from a highly virulent form of PRRSV. PRRSV nsp2 has recently been shown to interact or be part of a complex with potentially 285 cellular proteins.[38] In this study the authors used an immuno-precipitation approach and separation of products by SDS-PAGE and protein identification using MS. Our approach differed slightly from this analysis, in that we used a label free quantitative proteomic approach to identify potential binding partners. Our analysis identified 91 proteins that we assigned as having a potential interaction with HPnsp2. Several proteins were common to both studies; CTBP2, SSRG, CD2AP, RBM4, 14-3-3G, RL7A. RS7, NUCL, RL18A, RL13, 14-3-3E and lower order hits (in our study) such as ADT2, RL9, RL8, ILF3, RS3, PABP1, RPN1, 14-3-3T, HSP72, THOC4, HSP7C and HSP71. Binding

of nsp2 to a number of identified targets (NUCL, CD2AP, HSP70, DNAJA2, 14-3-3 and RS6) was confirmed using forward and reverse pull downs in separate repeat experiments (Figure 2) and immunofluorescence (Figure 3), placing confidence in the label free approach to correctly identify potential protein-protein interactions with HPnsp2.

14-3-3 is a highly conserved, ubiquitously expressed protein family. 14-3-3 proteins are involved in important cellular processes such as signal transduction, cell cycle control, apoptosis and the stress response by binding at least 100 different proteins.[39] Two phosphorylated consensus sequences were identified as 14-3-3 protein binding motifs[40], although these motifs are not found in nsp2. However, a non-consensus motif RX1-2SX2-3S (x represents any amino acid) was identified in 14-3-3 binding partner of Cbl[41] and can also be found in nsp2.. 14-3-3 may associate with the HV domain of nsp2 as indicated (Figure 9). 14-3-3 may have multiple roles[42] including as an adaptor molecule to facilitate protein-protein interactions. We postulate that 14-3-3 can act as an adaptor with nsp2. CD2AP was an adaptor protein which belongs to the CIN85/CD2AP and can act as a hub proteins[43] and play an integral role in actin remodeling.[44] The binding of CD2AP and nsp2 protein may change the structure of actin which facilitated the formation of aggresomes.

Our data indicated that nsp2 could associate with a number of cellular proteins involved in the formation of aggresomes. In virus-infected cells these have been associated with sites of virus replication and/or assembly.[33] This was investigated using a number of different complementary approaches included immunofluorescence (e.g. Figure 5) and biochemical analysis (e.g.

Figure 6). Both approaches indicated that nsp2 was associated with aggresomes.

Ribosomal proteins were also found to potentially interact with nsp2 and this may point towards a role of nsp2 in the modulation of this cellular activity and perhaps promotion of translation of viral gene products. Not all protein-protein interactions that nsp2 could form with cellular proteins were investigated and the proteomic approach provides a wealth of data that can be further investigated to probe PRRSV host cell interactions and the role of nsp2 in virus biology. For example, the data indicated that HPnsp2 formed an interaction with CTBP2 (C-terminal-binding protein 2). This is a transcriptional co-repressor that turns target genes off and is a potential target for viral intervention. Transcription of the mRNA encoding this protein has recently been shown to be repressed by Epstein-Barr virus (EBV) infected cells by the EBNA 3 protein.[45] Control of sub-cellular localisation is also emerging as an important mechanism whereby CTBP2 function is regulated. CTBP2 binds to client proteins through a PXDLS amino acid motif on the client. Analysis of the nsp2 amino acid sequence for both HPnsp2 and LPnsp2 and other strain variants indicates a conserved PXDLS motif, in this case PLDLS.

Analysis of the primary amino acid sequence of HPnsp suggested that it could potentially act as a hub protein with the HV region having the most disordered structure to facilitate protein-protein interactions (Figure 7).[46] The data indicated that CD2AP and 14-3-3 could form a protein-protein interaction with this region, but not the other regions that constitute HPnsp2. All seven isoforms of 14-3-3 were found to interact with HPnsp2 in our study.

In conclusion, the interactome analysis suggested that PRRSV nsp2 formed multiple cellular protein interactions that may have multiple implications for both viral and host cell biology. Targeting the function of these cellular proteins that may be critical for virus biology, with small molecular inhibitors, can be used as potential anti-viral therapy.

Acknowledgements

We would like to thank Dr. Ying Fang, College of Veterinary Medicine, Kansas State University, Manhattan KS, USA for providing us with pSD01-08-GFP plasmid.

References

(1) Wensvoort, G.; Terpstra, C.; Pol, J. M.; ter Laak, E. A.; Bloemraad, M.; de Kluyver, E. P.; Kragten, C.; van Buiten, L.; den Besten, A.; Wagenaar, F.; Broekhuijsen, J. M.; Moonen, P. L. J. M.; Zetstra T.; de Boer, E. A.; Tibben H. J.; de Jong, M. F.; van 't Veld, P.; Greenland, G. J. R.; van Gennep, J. A.; Voets, M.Th.; Verheijden J. H. M.; Braamskamp, J. Mystery swine disease in The Netherlands: the isolation of Lelystad virus. *Vet. Q.* **1991,** 13, 121-130.

(2) Collins, J. E.; Benfield, D. A.; Christianson, W. T.; Harris, L.; Hennings, J. C.; Shaw, D. P.; Goyal, S. M.; McCullough, S.; Morrison, R. B.; Joo, H. S.; Gorcyca, D.; Chladek, D. Isolation of swine infertility and respiratory syndrome virus (isolate ATCC VR-2332) in North America and experimental reproduction of the disease in gnotobiotic pigs. *J. Vet. Diagn. Invest.* **1992,** 4, 117-126.

(3) Wensvoort, G. Lelystad virus and the porcine epidemic abortion and respiratory syndrome. *Vet. Res.* **1993,** 24, 117-124.

(4) Cavanagh, D. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. *Arch. Virol.* **1997,** 142, 629-633.

(5) Wang, L.; Hou, J.; Zhang, H.; Feng, W. H. Complete genome sequence of a novel highly pathogenic porcine reproductive and respiratory syndrome virus variant. *J. Virol.* **2012,** 86, 13121.

(6) Wenhui, L.; Zhongyan, W.; Guanqun, Z.; Zhili, L.; JingYun, M.; Qingmei, X.; Baoli, S.; Yingzuo, B., Complete genome sequence of a novel variant porcine reproductive and respiratory syndrome virus (PRRSV) strain:

evidence for recombination between vaccine and wild-type PRRSV strains. *J Virol* **2012,** 86, 9543.

(7) Liu, D.; Zhou, R.; Zhang, J.; Zhou, L.; Jiang, Q.; Guo, X.; Ge, X.; Yang, H. Recombination analyses between two strains of porcine reproductive and respiratory syndrome virus in vivo. *Virus. Res.* **2011,** 155, 473-486.

(8) Tong, G. Z.; Zhou, Y. J.; Hao, X. F.; Tian, Z. J.; An, T. Q.; Qiu, H. J. Highly pathogenic porcine reproductive and respiratory syndrome, China. *Emerg. Infect. Dis.* **2007,** 13, 1434-1436.

(9) Zhou, L.; Yang, H. Porcine reproductive and respiratory syndrome in China. *Virus. Res.* **2010,** 154, 31-37.

(10) Snijder, E. J.; Meulenberg, J. J. The molecular biology of arteriviruses. *J. Gen. Virol.* **1 998,** 79, 961-979.

(11) Allende, R.; Lewis, T. L.; Lu, Z.; Rock, D. L.; Kutish, G. F.; Ali, A.; Doster, A. R.; Osorio, F. A. North American and European porcine reproductive and respiratory syndrome viruses differ in non-structural protein coding regions. *J. Gen. Virol.* **1999,** 80, 307-315.

(12) Nelsen, C. J.; Murtaugh, M. P.; Faaberg, K. S. Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. *J. Virol.* **1999,** 73, 270-280.

(13) Fang, Y.; Snijder, E. J. The PRRSV replicase: exploring the multifunctionality of an intriguing set of nonstructural proteins. *Virus. Res.* **2010,** 154, 61-76.

(14) Ziebuhr, J.; Snijder, E. J.; Gorbalenya, A. E. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* **2000,** 81, 853-879.

(15) Han, J.; Rutherford, M. S.; Faaberg, K. S. The porcine reproductive and respiratory syndrome virus nsp2 cysteine protease domain possesses both trans- and cis-cleavage activities. *J. Virol.* **2009,** 83, 9449-9463.

(16) Snijder, E. J.; Wassenaar, A. L.; Spaan, W. J.; Gorbalenya, A. E. The arterivirus Nsp2 protease. An unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases. *J. Biol. Chem.* **1995,** 270, 16671-16676.

(17) Wassenaar, A. L.; Spaan, W. J.; Gorbalenya, A. E.; Snijder, E. J. Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease. *J. Virol.* **1997,** 71, 9313-9322.

(18) Frias-Staheli, N.; Giannakopoulos, N. V.; Kikkert, M.; Taylor, S. L.; Bridgen, A.; Paragas, J.; Richt, J. A.; Rowland, R. R.; Schmaljohn, C. S.; Lenschow, D. J.; Snijder, E. J.; Garcia-Sastre, A.; Virgin, H. W. t. Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses. *Cell Host Microbe* **2007,** 2, 404-416.

(19) Sun, Z.; Li, Y.; Ransburgh, R.; Snijder, E. J.; Fang, Y. Nonstructural protein 2 of porcine reproductive and respiratory syndrome virus inhibits the antiviral function of interferon-stimulated gene 15. *J. Virol.* **2012,** 86, 3839-3850.

(20) Han, J.; Liu, G.; Wang, Y.; Faaberg, K. S. Identification of nonessential regions of the nsp2 replicase protein of porcine reproductive and respiratory syndrome virus strain VR-2332 for replication in cell culture. *J. Virol.* **2007,** 81, 9878-9890.

(21) Han, J.; Rutherford, M. S.; Faaberg, K. S. Proteolytic products of the porcine reproductive and respiratory syndrome virus nsp2 replicase protein. *J. Virol.* **2010,** 84, 10102-10112.

(22) Kappes, M. A.; Miller, C. L.; Faaberg, K. S. Highly divergent strains of porcine reproductive and respiratory syndrome virus incorporate multiple isoforms of nonstructural protein 2 into virions. *J. Virol.* **2013,** 87, 13456-13465.

(23) Chen, Z.; Zhou, X.; Lunney, J. K.; Lawson, S.; Sun, Z.; Brown, E.; Christopher-Hennings, J.; Knudsen, D.; Nelson, E.; Fang, Y. Immunodominant epitopes in nsp2 of porcine reproductive and respiratory syndrome virus are dispensable for replication, but play an important role in modulation of the host immune response. *J. Gen. Virol.* **2010,** 91, 1047-1057.

(24) Reed, L. J.; Muench, H. A. Simple method of estimating fifty percent endpoints. *Am. J. Hyg.* **1938,** 27, 493-497.

(25) Fang, Y.; Rowland, R. R.; Roof, M.; Lunney, J. K.; Christopher-Hennings, J.; Nelson, E. A. A full-length cDNA infectious clone of North American type 1 porcine reproductive and respiratory syndrome virus: expression of green fluorescent protein in the Nsp2 region. *J. Virol*. **2006,** 80,11447-11455.

(26) Wu, W.; Tran, K. C.; Teng, M. N.; Heesom, K. J.; Matthews, D. A.; Barr, J. N.; Hiscox, J. A. The interactome of the human respiratory syncytial virus NS1 protein highlights multiple effects on host cell biology. *J. Virol.* **2012,** 86, 7777-7789.

(27) Munday, D. C.; Wu, W.; Smith, N.; Fix, J.; Noton, S. L.; Galloux, M.; Touzelet, O.; Armstrong, S. D.; Dawson, J. M.; Aljabr, W.; Easton, A. J.; Rameix-Welti, M. A.; de Oliveira, A. P.; Simabuco, F.; Ventura, A. M.; Hughes, D. J.; Barr, J. N.; Fearns, R.; Digard, P.; Eleouet, J. F.; Hiscox, J. A. Interactome analysis of the human respiratory syncytial virus RNA polymerase complex identifies protein chaperones as important co-factors that promote L protein stability and RNA synthesis. J. Virol. **2014**, 89, 917-930.

(28) Emmott, E.; Munday, D.; Bickerton, E.; Britton, P.; Rodgers, M. A.; Whitehouse, A.; Zhou, E. M.; Hiscox, J. A. The cellular interactome of the coronavirus infectious bronchitis virus nucleocapsid protein and functional implications for virus biology. *J. Virol.* **2013,** 87, 9486-9500.

(29) Jourdan, S. S.; Osorio, F.; Hiscox, J. A. An interactome map of the nucleocapsid protein from a highly pathogenic North American porcine reproductive and respiratory syndrome virus strain generated using SILAC-based quantitative proteomics. *Proteomics* **2012,** 12, 1015-1023.

(30) Garcia-Dorival, I.; Wu, W.; Dowall, S.; Armstrong, S.; Touzelet, O.; Wastling, J.; Barr, J. N.; Matthews, D.; Carroll, M.; Hewson, R.; Hiscox, J. A. Elucidation of the Ebola virus VP24 cellular interactome and disruption of virus biology through targeted inhibition of host cell protein function. J. Proteome Res. **2014**, 13, 5120-5135.

(31) Kopito, R. R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* **2000,** 10, 524-530.

(32) Johnston, J. A.; Ward, C. L.; Kopito, R. R. Aggresomes: a cellular response to misfolded proteins. *J. Cell Biol.* **1998,** 143, 1883-1898.

(33) Wileman, T. Aggresomes and pericentriolar sites of virus assembly: cellular defense or viral design? *Annu. Rev. Microbiol.* **2007,** 61, 149-167.

(34) Lin, S. R.; Jiang, M. J.; Wang, H. H.; Hu, C. H.; Hsu, M. S.; Hsi, E.; Duh, C. Y.; Wang, S. K. Human cytomegalovirus UL76 elicits novel aggresome formation via interaction with S5a of the ubiquitin proteasome system. *J. Virol.* **2013,** 87, 11562-11578.

(35) Shen, D.; Coleman, J.; Chan, E.; Nicholson, T. P.; Dai, L.; Sheppard, P. W.; Patton, W. F. Novel cell- and tissue-based assays for detecting misfolded and aggregated protein accumulation within aggresomes and inclusion bodies. *Cell Biochem. Biophys.* **2011,** 60, 173-185.

(36) Xu, Z.; Graham, K.; Foote, M.; Liang, F.; Rizkallah, R.; Hurt, M.; Wang, Y.; Wu, Y.; Zhou, Y. 14-3-3 protein targets misfolded chaperone-associated proteins to aggresomes. *J. Cell Sci.* **2013,** 126, 4173-4186.

(37) Garcia-Mata, R.; Bebok, Z.; Sorscher, E. J.; Sztul, E. S. Characterization and dynamics of aggresome formation by a cytosolic GFP-chimera. *J. Cell Biol.* **1999,** 146, 1239-1254.

(38) Wang, L.; Zhou, L.; Zhang, H.; Li, Y.; Ge, X.; Guo, X.; Yu, K.; Yang, H. Interactome profile of the host cellular proteins and the nonstructural protein 2 of porcine reproductive and respiratory syndrome virus. *PLoS One* **2014,** 9, e99176.

(39) Xiao, B.; Smerdon, S. J.; Jones, D. H.; Dodson, G. G.; Soneji, Y., Aitken, A.; Gamblin, S. J. Structure of a 14-3-3 protein and implications for coordination of multiple signalling pathways. *Nature* **1995,** 376, 188-191.

(40) Dougherty, M. K.; Morrison, D. K. Unlocking the code of 14-3-3. *J. Cell Sci.* **2004,** 117, 1875-1884.

(41) Liu, Y. C.; Liu, Y.; Elly, C.; Yoshida, H.; Lipkowitz, S.; Altman, A. Serine phosphorylation of Cbl induced by phorbol ester enhances its association with 14-3-3 proteins in T cells via a novel serine-rich 14-3-3-binding motif. *J. Biol. Chem.* **1997,** 272, 9979-9985.

(42) van Hemert, M. J.; Steensma, H. Y.; van Heusden, G. P. 14-3-3 proteins: key regulators of cell division, signalling and apoptosis. *Bioessays*. **2001,** 23, 936-946.

(43) Rouka, E.; Simister, P. C.; Janning, M.; Kumbrink, J.; Konstantinou, T.; Muniz, J. R.; Joshi, D.; O'Reilly, N4.; Volkmer, R.; Ritter, B.; Knapp, S.; von Delft, F.; Kirsch, K. H.; Feller, S. M.Differential recognition preferences of the three Src homology 3 (SH3) domains from the adaptor CD2-associated protein (CD2AP), and direct association with Ras and Rab interactor 3 (RIN3). *J Biol Chem*. **2015,** Aug 20. pii: jbc.M115.637207. [Epub ahead of print].

(44) Gaidos, G.; Soni, S.; Oswald, D. J.; Toselli, P. A.; Kirsch, K. H. Structure and function analysis of the CMS/CIN85 protein family identifies actin-bundling properties and heterotypic-complex formation. *J Cell Sci*. **2007,** 120, 2366-2377.

(45) McClellan, M. J.; Wood, C. D.; Ojeniyi, O.; Cooper, T. J.; Kanhere, A.; Arvey, A.; Webb, H. M.; Palermo, R. D.; Harth-Hertle, M. L.; Kempkes, B.; Jenner, R. G.; West, M. J. Modulation of enhancer looping and differential gene targeting by Epstein-Barr virus transcription factors directs cellular reprogramming. *PLoS Pathog.* **2013,** 9, e1003636.

(46) Emmott, E.; Hiscox, J. A. Nucleolar targeting: the hub of the matter. *EMBO Rep.* **2009,** 10, 231-238.

Table 1 Cellular proteins showing more than 2 fold changes in abundance compared GFP-nsp2 with GFP

Protein identifier	Protein name	Peptides	Max fold change	Anova (p)	Protein function(s)	Molecular weight	Standard error (Control)	Standard error (SD16)
Nsp2	Nsp2	155	42.35	0.012	Nsp2_SD16	126242	456142.4	82564079.2
P56545	CTBP2	6	16.79	0.033	C-terminal binding protein 2; Corepressor targeting diverse transcription regulators.	48945	6730.8	88111.9
P62753	RS6	3	13.5	0.026	40S ribosomal protein S6; controlling cell growth and proliferation.	28681	2749.5	92520.5
P49207	RL34	3	11.88	0.001	60S ribosomal protein L34	13293	49.2	6375.3
Q9UNL2	SSRG	2	11.68	0.013	Signal sequence receptor, gamma (translocon-associated protein gamma); part of a complex of binding calcium to the ER membrane and retention of ER resident proteins	21080	245.5	4574.3
Q9Y5K6	CD2AP	12	11.68	0.004	CD2-associated protein; an adapter protein between membrane proteins and the actin;	71451	9697.0	6601.3

					connect receptor clustering and cytoskeletal polarity in the between T-cell and antigen-presenting cell.			
P61254	RL26	5	9.09	0.001	60S ribosomal protein L26	17258	1394.1	1640.1
P42766	RL35	3	8.85	0.029	60S ribosomal protein L35	14551	8105.0	8288.6
P62899	RL31	5	8.66	0.013	60S ribosomal protein L31	14463	10338.4	83616.6
O00264	PGRC1	2	7.7	0.118	Progesterone receptor membrane component 1; Receptor for progesterone	21671	3800.2	4434.1
Q9BWF3	RBM4	2	7.62	0.065	RNA binding motif protein 4; select alternative splice site during pre-mRNA processing	40314	540.0	16293.1
P98179	RBM3	3	7.28	0.029	RNA binding motif (RNP1, RRM) protein 3	17170	2347.0	571.5
P62910	RL32	4	7.22	0.008	60S ribosomal protein L32	15860	7801.5	14830.9
P31946	1433B	5	6.7	0.017	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide; Adapter protein in the regulation of a large spectrum of both general and specialized signaling pathway;	28082	4386.0	25272.1

					modulates the activity of the binding partner, usually by recognition of a phosphoserine or phosphothreonine motif.			
P62750	RL23A	2	6.39	0.06	60S ribosomal protein L23a	17695	3457.9	11014.1
P61981	1433G	6	6.17	0	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide; similar function with 1433B	28303	365.5	1784.7
P61353	RL27	2	6.14	0.026	60S ribosomal protein L27	15798	6713.3	90783.5
P62424	RL7A	4	5.63	0.003	60S ribosomal protein L7a	29996	1349.9	9809.4
P62081	RS7	9	5.61	0.02	40S ribosomal protein S7; Required for rRNA maturation	22127	34601.3	309587.4
P19338	NUCL	23	5.56	0.014	Nucleolin; induces chromatin decondensation; facilitates pre-rRNA transcription and ribosome assembly and transcriptional elongation.	76614	136237.1	1163450.8
P46779	RL28	11	5.52	0.004	60S ribosomal protein L28	15748	3223.1	59282.1
Q02543	RL18A	4	5.46	0.005	60S ribosomal protein L18a	20762	1642.6	2944.9

P26373	RL13	9	5.42	0.007	60S ribosomal protein L13	24261	4649.1	182954.7
P62263	RS14	4	5.4	0.03	40S ribosomal protein S14	16273	15012.6	104547.2
P51571	SSRD	2	5.23	0.072	Translocon-associated protein subunit delta	18999	2252.5	12540.3
P82673	RT35	2	5.17	0.046	28S mitochondrial ribosomal protein S35	36844	1887.9	6548.9
Q02878	RL6	3	5.14	0.033	60S ribosomal protein L6	32728	1523.0	34314.6
O60884	DNJA2	2	4.71	0.063	DnaJ (Hsp40) homolog, subfamily A, member 2; Co-chaperone of Hsc70	45746	3981.5	17105.1
P84103	SRSF3	2	4.56	0.039	Splicing factor, arginine/serine-rich 3; involved in RNA processing in relation with cellular proliferation and/or maturation	19330	4737.1	33604.1
P62847	RS24	2	4.5	0.165	40S ribosomal protein S24; subunit of processing of pre-rRNA and maturation of 40S	15423	1857.7	18473.8
P46778	RL21	12	4.49	0	60S ribosomal protein L21	18565	699.9	28848.7
P62258	1433E	6	4.4	0.003	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide; similar function with	29174	2550.1	84910.6

					1433B			
P46783	RS10	4	4.15	0.013	40S ribosomal protein S10	18898	5928.9	18457.0
P62266	RS23	3	4	0.004	40S ribosomal protein S23	15808	421.4	2830.6
P61513	RL37A	5	3.74	0.103	60S ribosomal protein L37a	10275	8014.2	16382.5
P62273	RS29	2	3.55	0.021	40S ribosomal protein S29	6677	8117.6	10328.9
P47914	RL29	5	3.52	0.018	60S ribosomal protein L29	17752	402.3	27449.9
P60866	RS20	4	3.49	0.006	40S ribosomal protein S20	13373	5813.6	17178.3
Q9NZI8	IF2B1	2	3.44	0	Insulin-like growth factor 2 mRNA binding protein 1; RNA-binding factor; stable mRNA	63481	5.7	249.5
Q9Y3U8	RL36	2	3.43	0.002	60S ribosomal protein L36	12254	1265.1	2049.9
Q9NX63	CHCH3	2	2.98	0.285	Coiled-coil-helix-coiled-coil-helix domain containing 3	26152	4577.7	7606.7
P05141	ADT2	4	2.93	0.014	ADP/ATP translocase 2, catalyzes the exchange of ADP and ATP across the mitochondrial inner membrane	32852	5358.2	71831.3

P62857	RS28	8	2.92	0.029	40S ribosomal protein S28	7841	9829.9	63533.9
Q99623	PHB2	2	2.89	0.069	Prohibitin 2; mediator of transcriptional repression, Coregulator of estrogen receptor (ER) and represse the activity of estrogens; involved in regulating mitochondrial respiration activity and in aging	33296	1116.6	10724.6
P62829	RL23	4	2.88	0.216	60S ribosomal protein L23	14865	10755.2	27857.4
P32969	RL9	5	2.84	0.049	60S ribosomal protein L9	21863	6206.3	48571.7
P39019	RS19	2	2.77	0.016	40S ribosomal protein S19; Required for pre-rRNA processing and maturation of 40S ribosomal subunits	16060	1808.1	3878.6
P13073	COX41	2	2.69	0.009	Cytochrome c oxidase subunit IV isoform 1; nuclear-coded polypeptide chains of cytochrome c oxidase, the terminal oxidase in mitochondrial electron transport	19577	204.4	3891.7
P62888	RL30	2	2.69	0.083	60S ribosomal protein L30	12784	5962.1	64807.6
P35268	RL22	3	2.68	0.117	60S ribosomal protein L22	14787	5598.8	189963.8

P62249	RS16	6	2.63	0.019	40S ribosomal protein S16	16445	1147.6	18901.4
Q8NC51	PAIRB	10	2.53	0.076	Plasminogen activator inhibitor 1 RNA-binding protein; regulation of mRNA stability, binds to the 3'-most 134 nt of the SERPINE1/PAI1 mRNA.	44965	39327.9	11552.6
P37108	SRP14	2	2.52	0.018	Signal recognition particle 14kDa (homologous Alu RNA binding protein); portion of the SRP RNA, required for SRP RNA binding	14570	1842.1	3482.2
P62913	RL11	4	2.47	0.006	60S ribosomal protein L11; Binds to 5S ribosomal RNA. Required for rRNA maturation and formation of the 60S ribosomal subunits	20252	69.0	3135.3
P62917	RL8	5	2.45	0.013	60S ribosomal protein L8	28025	1264.7	11119.6
Q04917	1433F	2	2.42	0.079	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide, similar function with 1433B	28219	12217.4	12829.8
P62280	RS11	12	2.39	0.011	40S ribosomal protein S11	18431	17724.4	2875.1

P83731	RL24	3	2.39	0.002	60S ribosomal protein L24	17779	98.3	67.5
P63220	RS21	4	2.38	0.103	40S ribosomal protein S21	9111	10021.0	5755.4
P0CW22	RS17L	6	2.23	0.108	40S ribosomal protein S17-like	15550	84630.2	348454.8
Q13347	EIF3I	2	2.19	0.056	Eukaryotic translation initiation factor 3, subunit I; component of the eukaryotic translation initiation factor 3 (eIF-3) complex for the initiation of protein synthesis.	36502	19.5	1413.3
Q5JNZ5	RS26L	2	2.18	0.004	40S ribosomal protein S26-like	13002	1095.2	1349.6
P63104	1433Z	5	2.16	0.105	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide; similar function with 1433B	27745	13152.6	25667.6
P20674	COX5A	2	2.14	0.075	Cytochrome c oxidase subunit Va; heme A-containing chain of cytochrome c oxidase, the terminal oxidase in mitochondrial electron transport	16762	2784.1	239.8
P25398	RS12	2	2.09	0.048	40S ribosomal protein S12	14515	1248.6	11021.0
P30101	PDIA3	2	2.06	0.22	Protein disulfide isomerase family, A	56782	4355.4	12153.3

member 3

Figure legends

Figure 1. Expression of EGFP-HPnsp2 and EGFP-LPnsp2 in 293T cells. The expression of EGFP fused nsp2 proteins were validated by fluorescence microscopy (A) and Western blot (B). The predicted molecular size was 26.9, 153,157 kDa for EGFP, EGFP-HPnsp2 and EGFP-LPnsp2, respectively.

Figure 2. Validation of the interaction of nsp2 with cellular proteins by Western blot. (A) Confirmation of the proteins associated with nsp2 in EGFP, EGFP-HPnsp2 and EGGP-LPnsp2 transfected 293T cells in GFP-trap pull-down using specific antibodies against 14-3-3, CD2AP, HSP70, DNAJA2, NUCL orRS6. A RFP-trap pull-down and non-transfected 293T cells were also used as the negative controls. (B) Confirmation of the proteins associated with nsp2 by reverse immunoprecipitation. EGFP, EGFP-HPnsp2 and EGFP-LPnsp2 transfected 293T cell lysates were incubated with specific antibodies against NUCL, CD2AP, HSP70, DNAJA2, 14-3-3 and RS6 after pre-clearing with blank protein G resin, respectively. The existence of EGFP-HPnsp2 or EGFP-LPnsp2 in the input samples and the eluates were detected by a specific antibody against GFP.

Figure 3. Co-localization of nsp2 with cellular proteins confirmed by immunofluorescence microscopy. The co-localization of EGFP-HPnsp2 and EGFP-LPnsp2 with 14-3-3, CD2AP or RS6 was visualized in both transfected 293T cells (used for the pull down), and MARC-145 cells (a permissive cell line for PRRSV). EGFP-HPnsp2 and EGFP-LPnsp2 were expressed in PK-15 cells (a porcine originated cell line) and 14-3-3 and CD2AP visualised. The

EGFP-tagged proteins were shown in green and cellular proteins were shown in red. The co-localization was determined by the yellow signal in the merged images.

Figure 4. Validation of the interaction of nsp2 with cellular proteins in PRRSV infected MARC-145 cells. (A) Western blot analysis of the input and bound samples by GFP-trap pull-down in MARC-145 cells either mock infected or infected with a European PRRSV-SD01-08 strain carrying a GFP reporter gene in nsp2 gene at a MOI of 3. MARC-145 cells transfected with EGFP-N1 plasmid were used as a control for the proteins binding to EGFP. All cells were harvested 30hpi. Proteins interacting with nsp2 were pulled down by GFP-Trap. The pulldown samples were separated by SDS-PAGE and detected by specific antibodies against GFP, HSP70, 14-3-3, NUCL and CD2AP. (B) Immunofluorescence microscopy showing the co-localization of 14-3-3 and viral proteins. MARC-145 cells were either mock infected or infected with both type 2 LP-PRRSV CH-1R strain and HP-PRRSV SD16 strain at an MOI 0.01. The cells were fixed at 24 hours post-transfection. PRRSV proteins were visualized with FITC-goat anti-pig IgG in green and 14-3-3 were visualized with Cy3-goat anti-rabbit IgG in red. The co-localization was determined by the yellow signal in the merged images.

Figure 5. Aggresome detection in EGFP, EGFP-HPnsp2 or EGFP-LPnsp2 transfected 293T and Marc-145 cells. (A) Co-localization of EGFP-HPnsp2 or EGFP-LPnsp2 with aggresomes maker proteins labelled by a novel 488 nm excitable red fluorescent molecular rotor dye which specifically detects

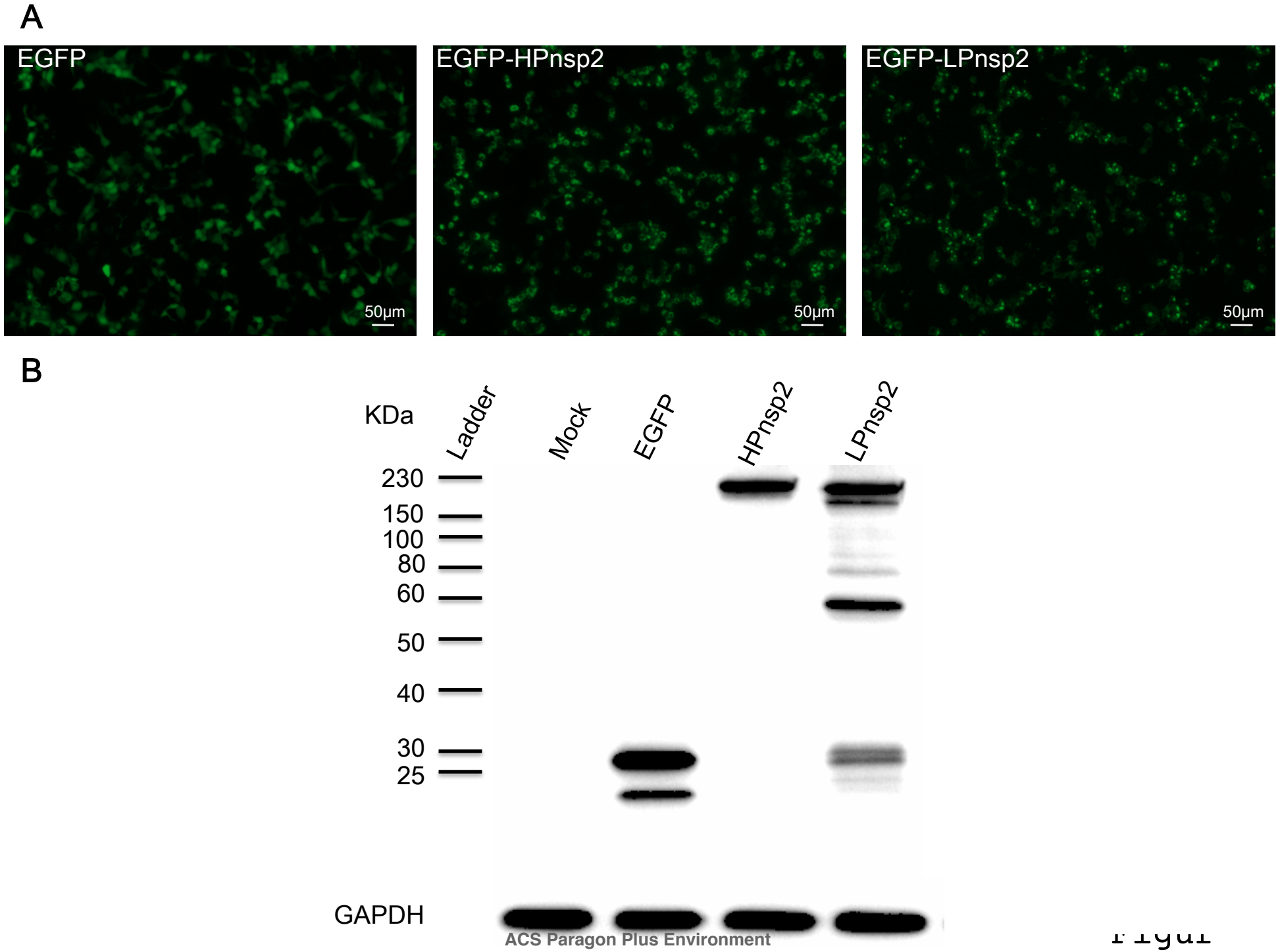
aggregated proteins within the aggresome and aggresome-like structures; (B) Co-localization of EGFP-HPnsp2 or EGFP-LPnsp2 with the aggresome component proteins HSP70 and Vimentin detected with anti-HSP70 and vimentin antibodies in red. The co-localization was determined by the yellow signal in the merged images.

Figure 6. Identification of cellular proteins involved in aggresome formation by Western blot. EGFP, EGFP-LPnsp2 and EGFP-HPnsp2 transfected 293T cells were collected and lysed 24 hours post transfection. (A) To separate cellular proteins that interacted with 14-3-3, the whole cell lysates were incubated with anti-14-3-3 antibody followed by protein G beads. The pull-down products were separated by SDS-PAGE and detected with specific antibodies against GFP, CD2AP, HSP70 and DNAJA2. (B) To enrich aggresomes, the whole cell lysates were centrifuged at high speed after brief centrifugation at low speed. The soluble fraction and insoluble fraction were collected. The insoluble fraction was collected after stringent wash with RIPA buffer and was solved in 8M urea buffer. The soluble and insoluble fractions were separated by SDS-PAGE and detected with specific antibodies against GFP, CD2AP, HSP70, DNAJA2, 14-3-3 and GAPDH.

Figure 7. Prediction of the disorder region of PRRSV HPnsp2 protein using an online PrDOS protein disorder prediction system at a prediction false positive rate of 5%.

Figure 8. Expression of truncated HPnsp2 proteins in 293T cells 24 hour post-transfection. (A) Schematic diagram of truncated nsp2 proteins with four concession fragments OTU, HV, TM, Tail spanning 47-159aa, 160-814aa, 815-988aa, and 989-1164aa, respectively. (B) Expression of truncated HPnsp2 proteins visualized by direct fluorescence microscopy. (C) Western blot analysis of the truncated nsp2 proteins expressed in 293T cells and detection with a specific anti-GFP antibody. EGFP, OTU-EGFP, TM-EGFP, Tail-EGFP, HPnsp2-EGFP with predicted molecular weights of 26.9kDa, 38.2kDa, 98.6kDa, 46kDa, 46.1kDa or 153.1kDa in order.

Figure 9. Identification of specific domains of HPnsp2 interacting with cellular protein targets by western blot. 293T cells transfected with plasmids expressing each truncated HPnsp2 as well as full length HPnsp2 and EGFP as controls were harvested and lysed 24 hours post transfection and subjected to GFP-trap pull-down. The cellular proteins interacted with each truncated HPnsp2 protein was detected by specific antibodies against NUCL, CD2AP, HSP70, DNAJA2, 14-3-3 or HSP90 in both input samples and in the eluates.


A
EGFP
50μm
EGFP-HPnsp2
50μm
EGFP-LPnsp2
50μm
B
KDa
Ladder
Mock
EGFP
HPnsp2
LPnsp2
230
150
100
80
60
50
40
30
25
GAPDH

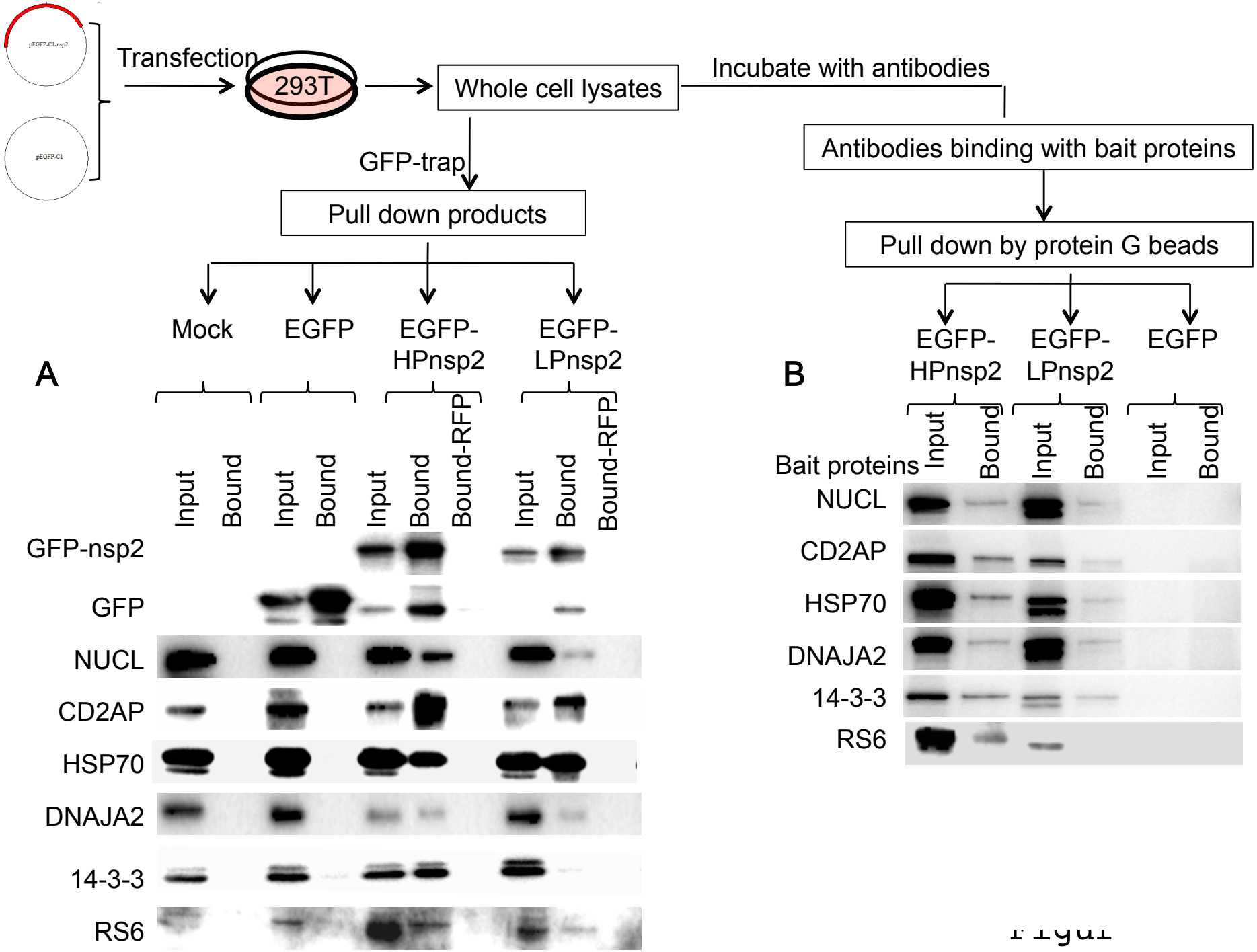
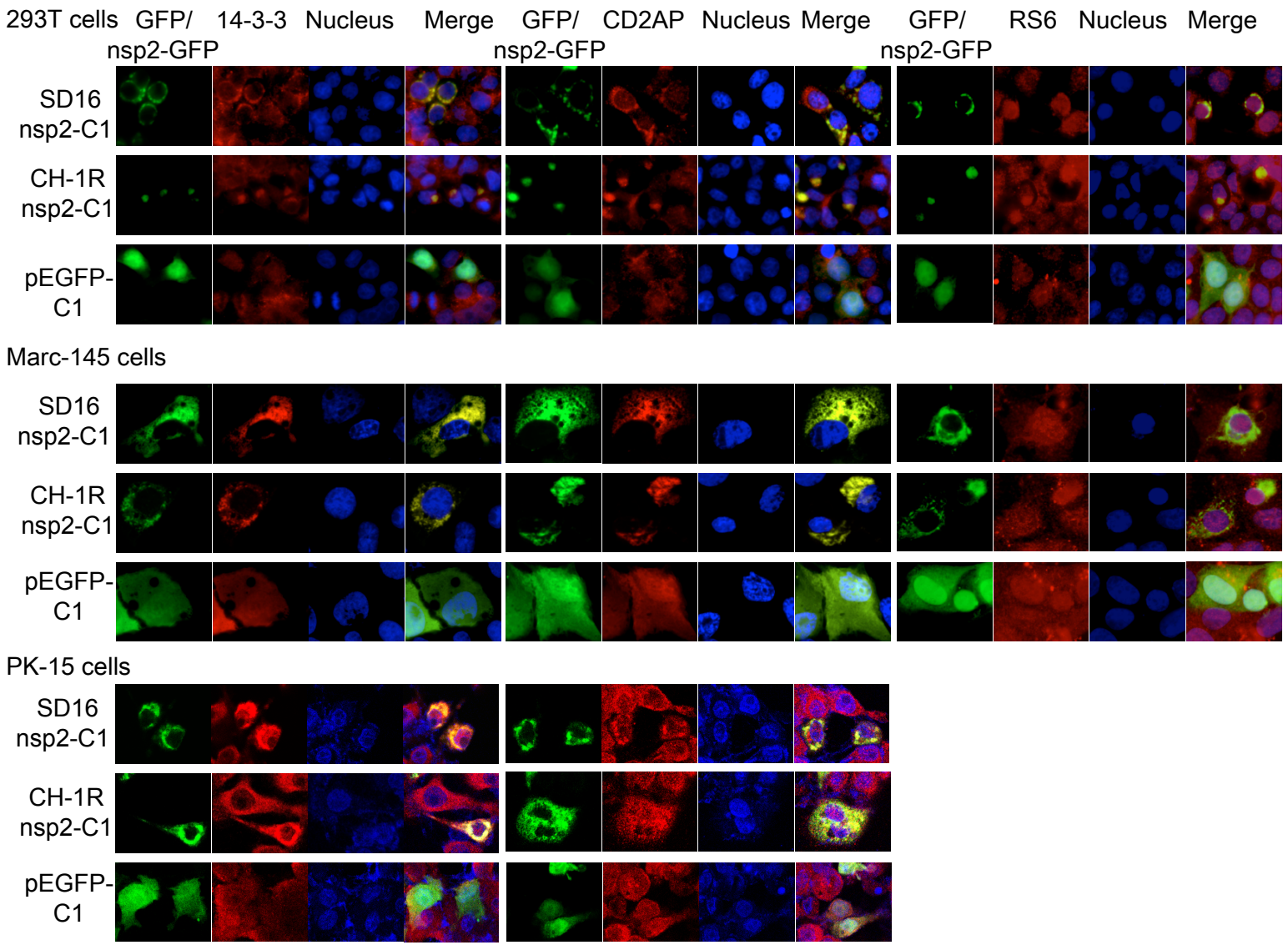

pEGFP-C1-nsp2
pEGFP-C1
Transfection
293T
Whole cell lysates
Incubate with antibodies
GFP-trap
Pull down products
Antibodies binding with bait proteins
Pull down by protein G beads
Mock
EGFP
EGFP-HPnsp2
EGFP-LPnsp2
A
B
Input
Bound
Bound-RFP
Bait proteins
GFP-nsp2
GFP
NUCL
CD2AP
HSP70
DNAJA2
14-3-3
RS6

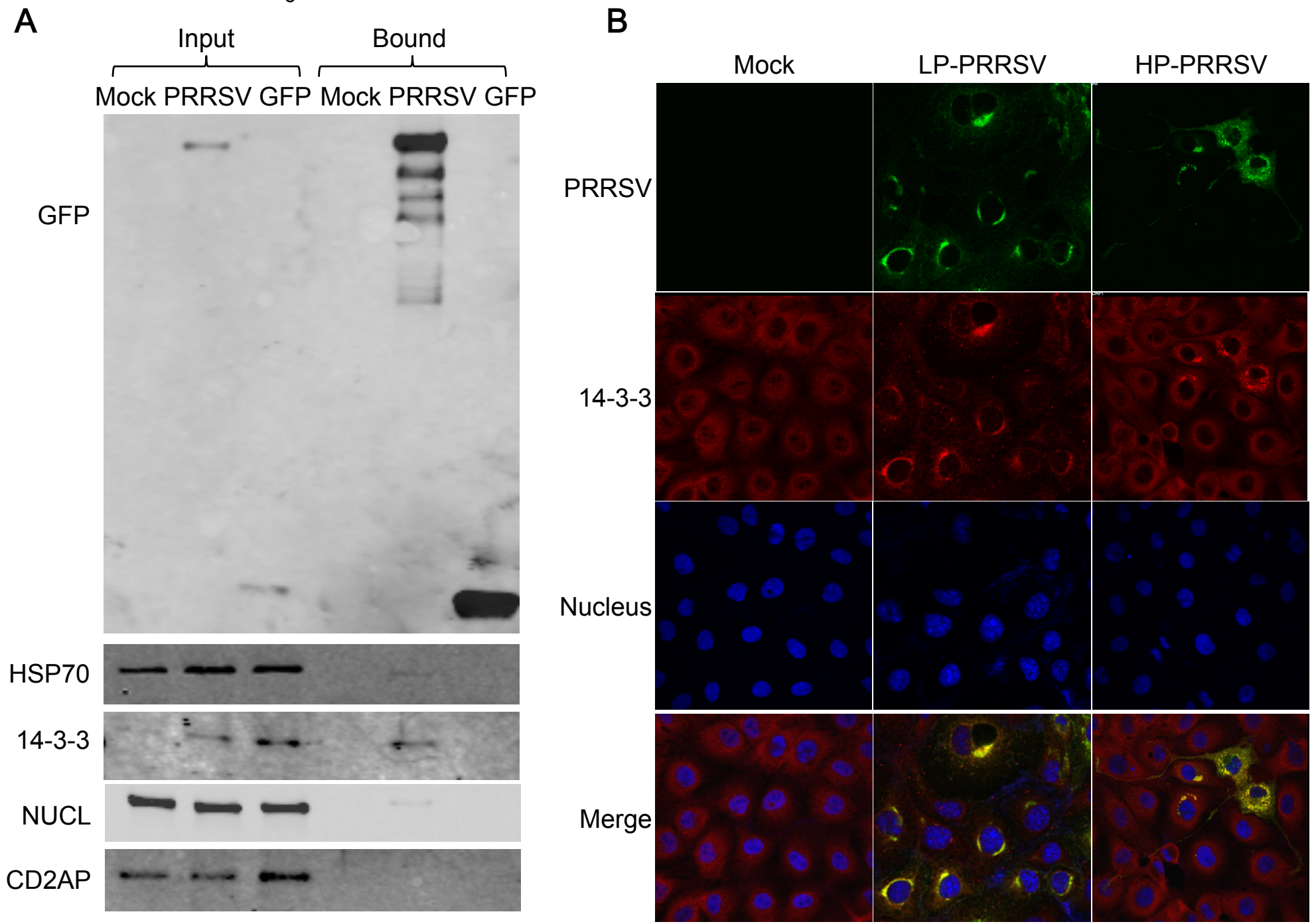
Figure 2


293T cells
GFP/ nsp2-GFP
14-3-3
Nucleus
Merge
GFP/ nsp2-GFP
CD2AP
Nucleus
Merge
GFP/ nsp2-GFP
RS6
Nucleus
Merge
SD16 nsp2-C1
CH-1R nsp2-C1
pEGFP-C1
Marc-145 cells
SD16 nsp2-C1
CH-1R nsp2-C1
pEGFP-C1
PK-15 cells
SD16 nsp2-C1
CH-1R nsp2-C1
pEGFP-C1

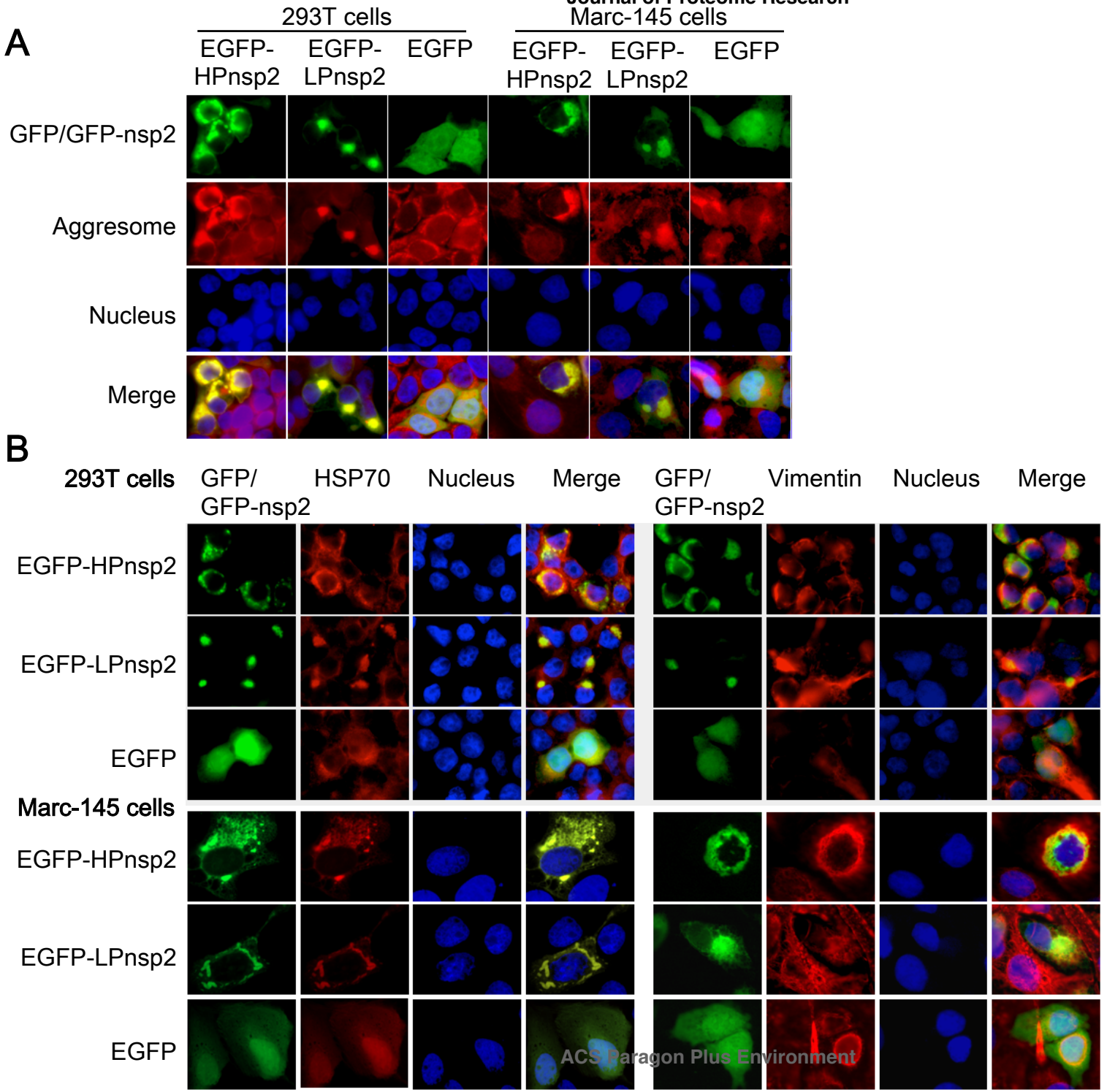
e

Figure 2

e


A
Input
Bound
Mock PRRSV GFP Mock PRRSV GFP
GFP
HSP70
14-3-3
NUCL
CD2AP
B
Mock
LP-PRRSV
HP-PRRSV
PRRSV
14-3-3
Nucleus
Merge

Figure

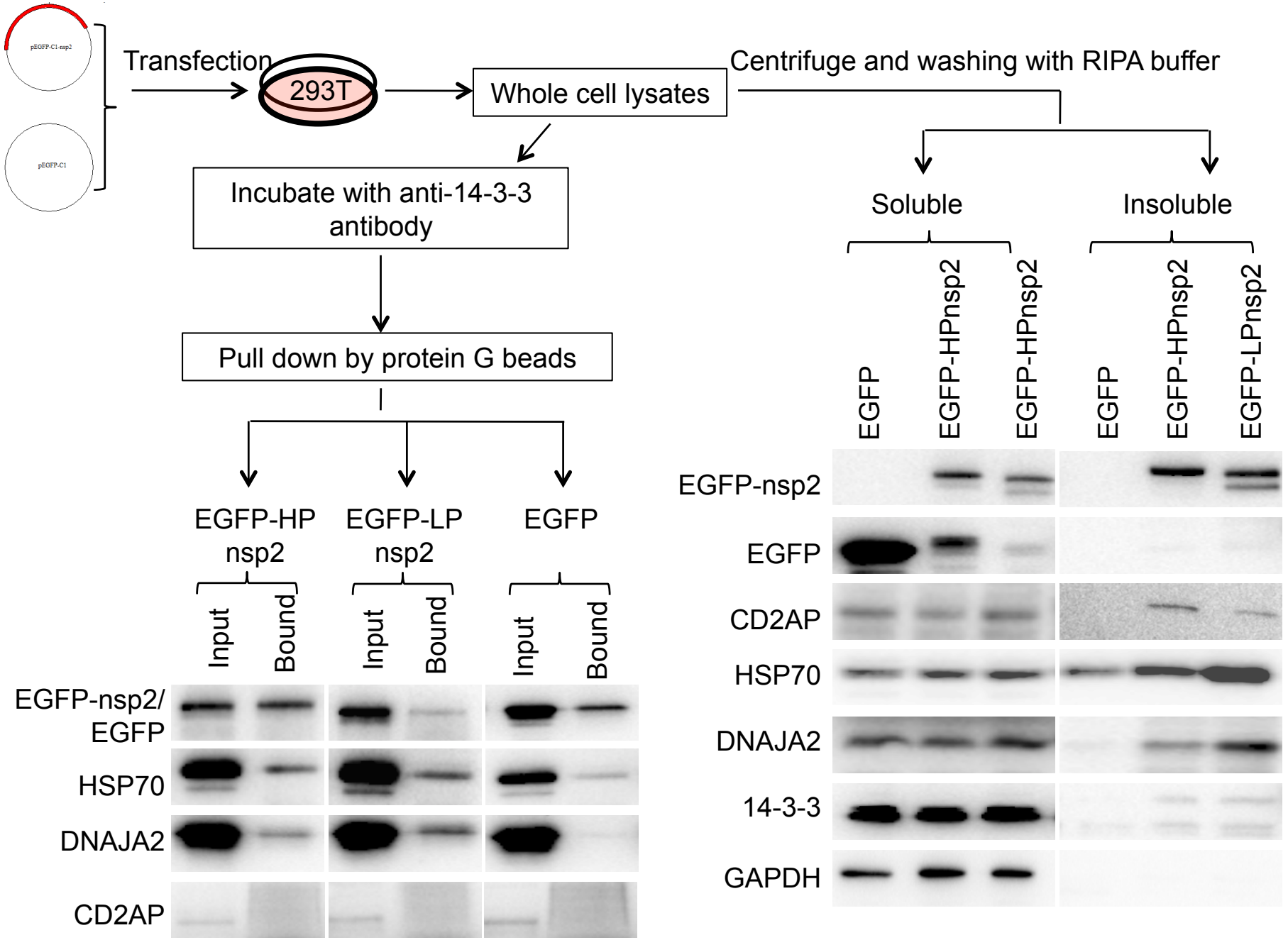


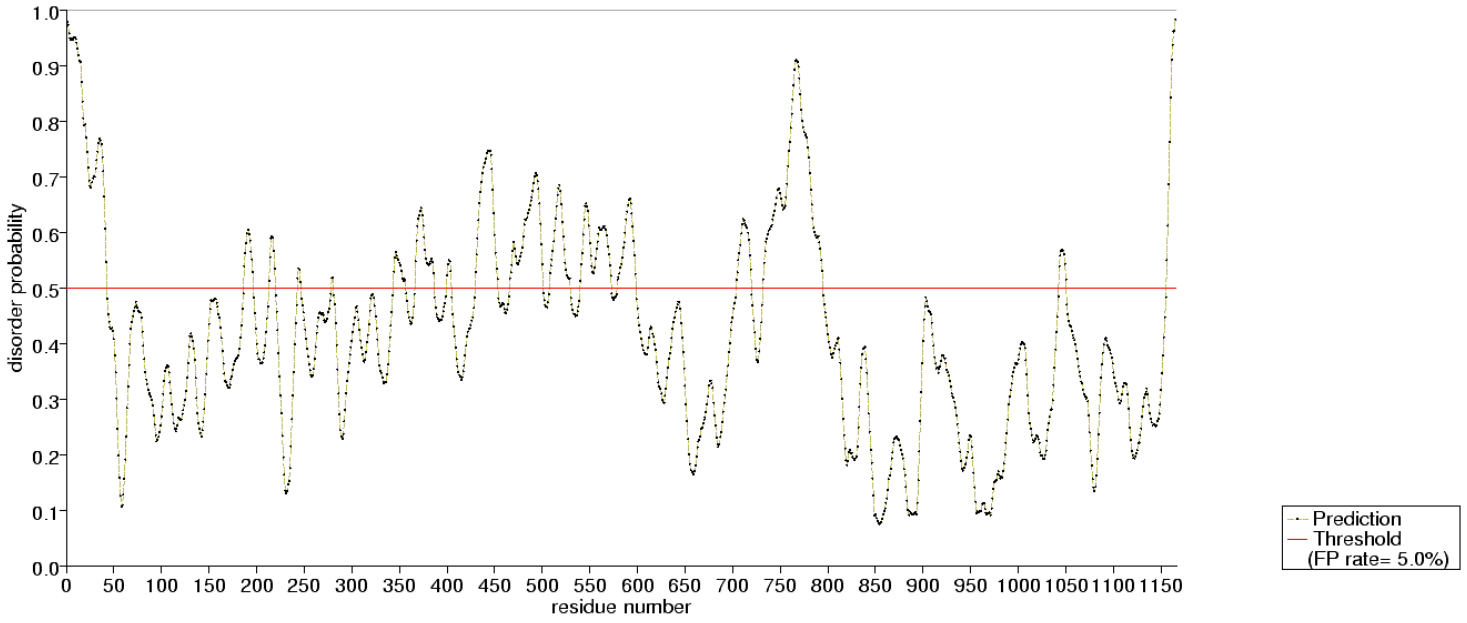
ACS Paragon Plus Environment

Figure

e

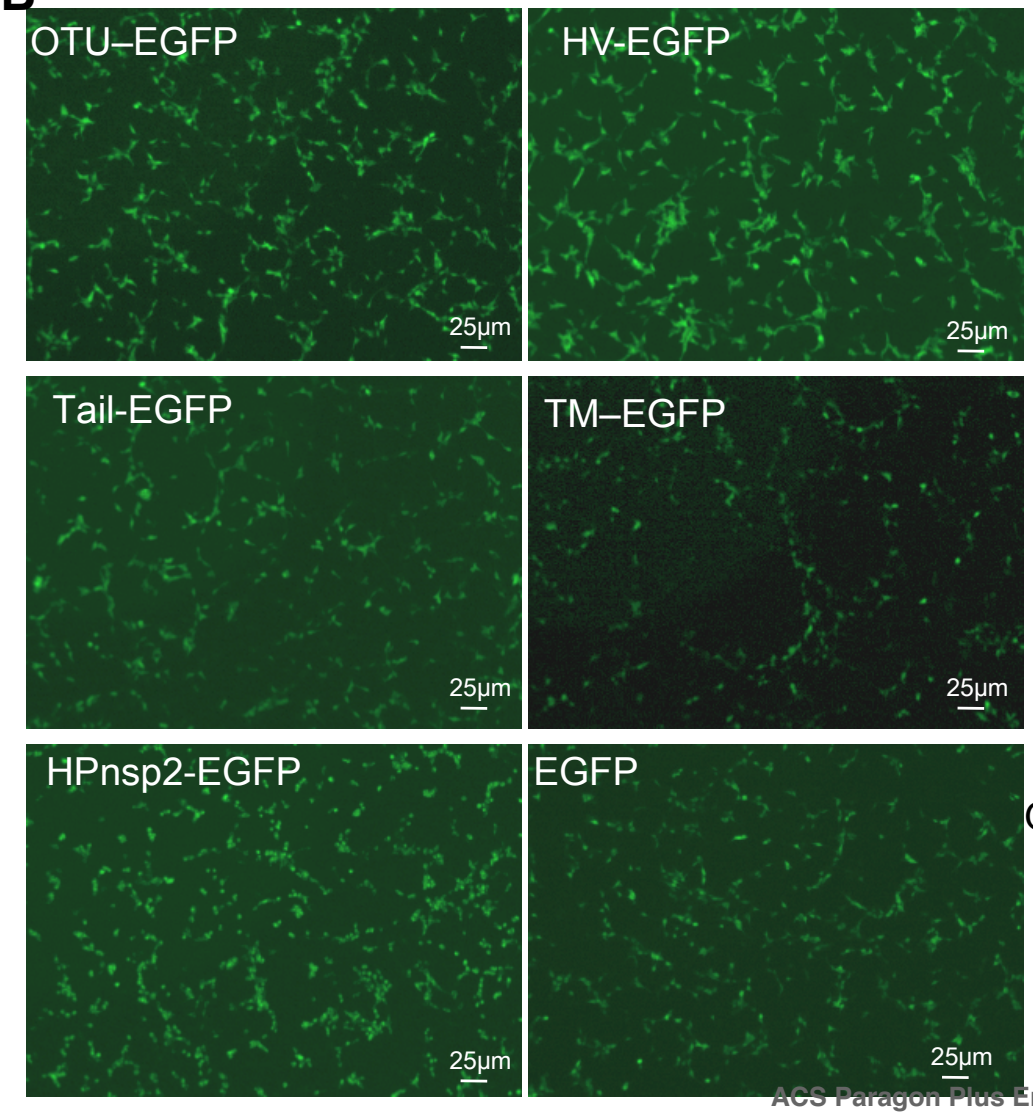
e

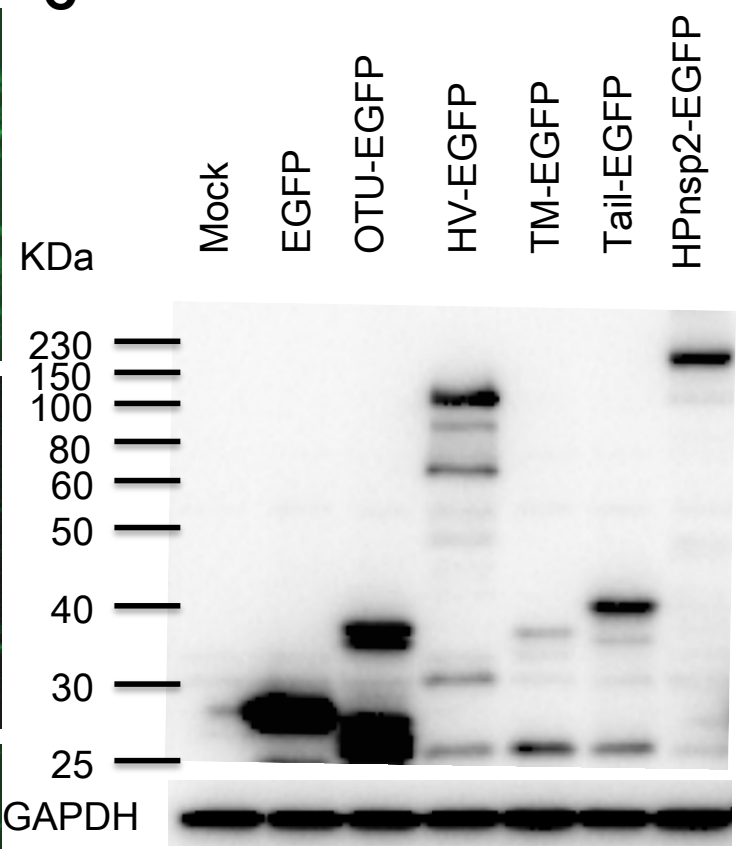

pEGFP-C1-nsp2
pEGFP-C1
Transfection
293T
Whole cell lysates
Centrifuge and washing with RIPA buffer
Incubate with anti-14-3-3 antibody
Pull down by protein G beads
Soluble
Insoluble
EGFP
EGFP-HPnsp2
EGFP-HPnsp2
EGFP
EGFP-HPnsp2
EGFP-LPnsp2
EGFP-nsp2
EGFP
CD2AP
HSP70
DNAJA2
14-3-3
GAPDH
EGFP-HP nsp2
EGFP-LP nsp2
EGFP
Input
Bound
Input
Bound
Input
Bound
EGFP-nsp2/ EGFP
HSP70
DNAJA2
CD2AP


disorder probability
residue number
Prediction
Threshold
(FP rate= 5.0%)

e

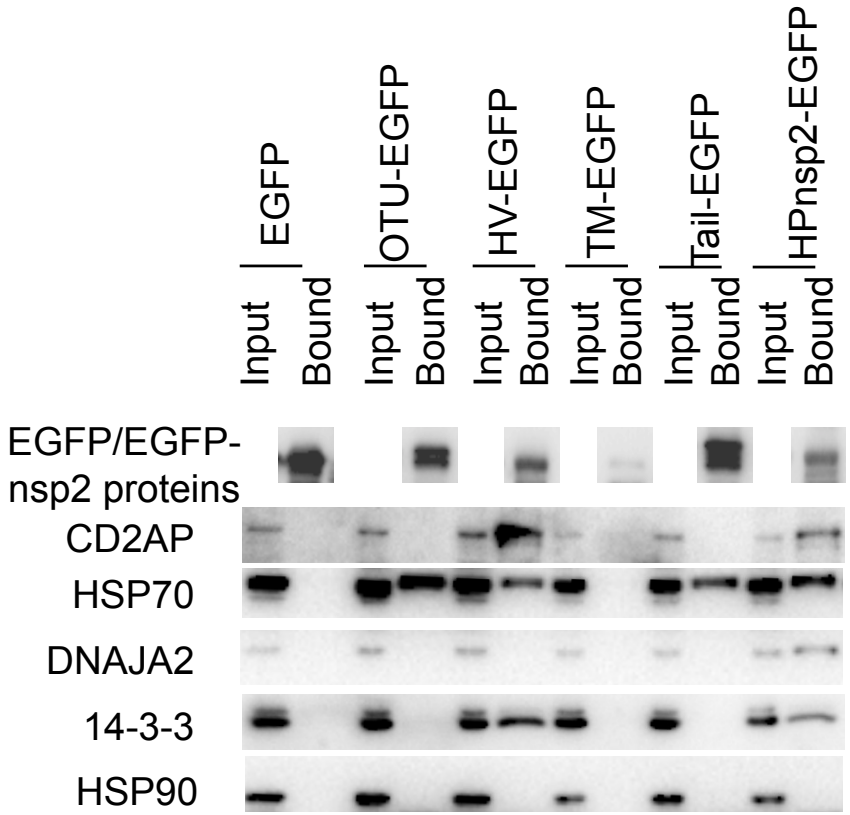

A
Nsp2(aa1-1194)
OUT(aa47-159)
HV(aa160-814)
TM(aa815-988)
Tail(aa989-1164)


B
OTU–EGFP
HV-EGFP
Tail-EGFP
TM–EGFP
HPnsp2-EGFP
EGFP
25μm


C
Mock
EGFP
OTU-EGFP
HV-EGFP
TM-EGFP
Tail-EGFP
HPnsp2-EGFP
KDa
230
150
100
80
60
50
40
30
25
GAPDH

Figure

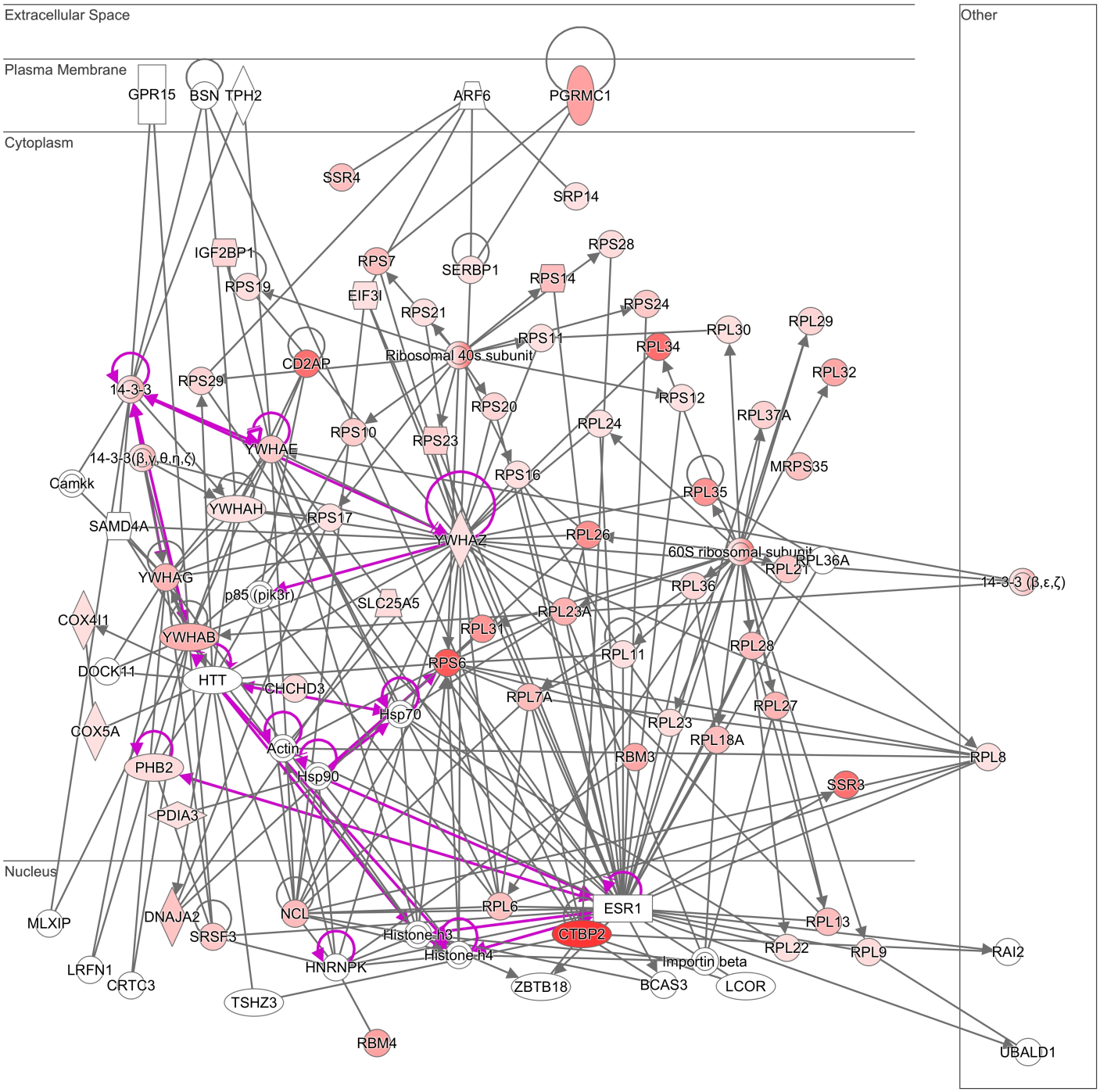
e


EGFP
OTU-EGFP
HV-EGFP
TM-EGFP
Tail-EGFP
HPnsp2-EGFP
Input
Bound
EGFP/EGFP-nsp2 proteins
CD2AP
HSP70
DNAJA2
14-3-3
HSP90

Figure

e

The interactome of porcine reproductive and respiratory syndrome virus nsp2



For TOC only